《西方将主宰多久》续篇

颠覆历史研究的惯常思维：文明可用数据来度量

从上万年的历史角度探寻东西方世界文明繁盛的根源

解决“大历史”观的问题：文明何以兴起？又何以衰落

震撼性的研究成果：未来世界的文明将由东方还是西方主导

文明的度量

社会发展如何决定国家命运

［美］伊恩·莫里斯（Ian Morris）◎著

李阳◎译

The Measure of Civilization

How Social Development Decides the Fate of Nations



中信出版社·CHINA CITIC PRESS

版权信息

书名:文明的度量：社会发展如何决定国家命运

作者:[美]伊恩·莫里斯

译者:李阳

ISBN:9787508644943

中信出版集团制作发行

The Measure of Civilization

How Social Development Decides the Fate of Nations

第一章

我们如何度量文明

为何西方主宰了世界

大约250年前，西欧的知识分子们遇到了一个问题。问题本身并不算坏：西方似乎主宰着世界，却不知道何以如此。18世纪的理论家们的解释千差万别，不过最流行的几种观点都认为：自远古时代起，西方就形成了一些与世界上其他地方不同的特质，这些特质决定了西方终有一天要成为世界霸主。

直到21世纪初，这样的观点仍然有市场，只是形式得到了极大的改良。其中最具影响力的观点与18世纪的如出一辙，是认为欧洲人继承了某种独特而先进的文化传统。这种西方文明的根源通常要追溯到古希腊和古罗马，不过也有人认为，史前印欧人、古日耳曼人，或者说中世纪的欧洲人，才是西方文明的奠基者。

在18世纪，还有一股思潮认为，正是得天独厚的环境和气候条件，使得欧洲人比其他人种精力更充沛、创造力更强。这种观点至今仍有大量支持者。一些学者将生态论和文化论结合起来，提出是两者的相互作用将现代化初期的欧洲推上了一条崭新的道路。甚至连欧洲人在生理上优于其他人种的观点，也在改头换面后仍为一些人所坚持：一些经济学家声称，自13世纪以来，自然选择使得欧洲人比其他任何人种都更节俭、更勤劳；与此同时，也有很多古人类学家认为，自一万年前农业起源以来，遗传进化方面的差异使得欧洲人及其后代比其他人种更具活力、更善于发明创造。

这些理论全都成形于18世纪，当时欧洲的财富暴涨和实力激增的现象迫切需要得到解释。直到20世纪晚期，在东亚也出现了同样的暴涨现象后，这些理论才遭遇严峻挑战。随着日本、“亚洲四小龙”和中国跃居主要经济体，越来越多的学者认识到，用长期的文化、环境或种族因素来解释西方成功的理论，肯定是不正确的。他们开始认为，

世界历史的重头戏并非西方长期而不可阻拦的崛起，而是多极世界的形成，西方只是在近代暂时甚至可能是偶然占据了优势。

这些新理论甚至比旧的长期注定论的争议还要大。其中一些最为极端的观点认为18世纪的理论家们的认识恰好是本末倒置的。按照这些新理论，实际上中国曾经在一个相当长的历史时期内在全球占据领先优势，只是一系列怪异的偶然事件才使得天平短暂地向欧洲倾斜。然而，绝大多数理论观点完全反对进行长期分析，认为复杂的亚洲社会和欧洲社会是沿着平行的轨道发展的，直到18世纪，甚至19世纪，由于双方在国家结构、先天条件、自然地理和政治地理，或者思想潮流方面的细小差异，才使欧洲获得了统治地位。

关于西方崛起的原因和后果的争论，引起了很多人的兴趣，但不同理论的支持者似乎经常是在各说各话。他们常常以不同的方式定义关键术语，使用不同类型的证据，并采用不同的举证标准。结果，观点不同者很难就他们试图说明什么达成一致，更不用提如何说明。

在我看来，真正需要讨论的问题，是我称之为“社会发展”的问题，也就是社会通过影响物理、经济、智力等自然环境和知识环境以达到相应目标的能力。新版本的18世纪理论的捍卫者们倾向于认为，西方的社会发展在数百年前，甚至数千年前就已经领先于世界其他地方了。他们的批评者则倾向于认为，西方只是在六七代人之前才脱颖而出的。如果真想解释西方称霸的原因，就需要衡量社会发展，对其进行跨越时空的比较。我们只有在建立了社会发展史的基本模式后，才能着手分析历史为什么是这样演进的。

定量分析并不一定使争论更加客观，但的确通常能使之更清晰，能促使争论各方讲清楚他们所使用的术语究竟是什么意思，阐明他们为什么要赋予这些差异不同的数值。任何不同意某位学者观点的人，都可以关注其证据和用于计算分值的方法，而不是交换含混不清、道理不足的概括总结。与社会发展概念类似的数值指数，以这样或那样

的名称，在人类学、考古学、经济学、金融学、社会学和政策制定等领域，都已经得到了完善，而且在联合国人类发展指数（Human Development Index，简称HDI）中，也有一个这样尺度鲜明的模式。

20世纪六七十年代，一些历史学家也开始将类似方法引入历史研究，通过搜集大量统计数据来解决大问题。其经典的案例也许当属罗伯特·福格尔（Robert Fogel）和斯坦利·恩格尔曼（Stanley Engerman）合著的《十字架上的时间》（*Time on the Cross*）。该书汇集了数千份种植园档案中的数据，分析出19世纪美国南方的奴隶制究竟怎样有利可图，以及奴隶们又有怎样的亲身经历。

《十字架上的时间》为计量史学提供了一个成功的模式。该书共有两卷，第一卷是较宽泛的概述，既针对专业学者，也面向对美国历史感兴趣的一般读者；第二卷则详尽地展现了福格尔和恩格尔曼的统计技术和对统计资料的运用。

《文明的度量》将遵循这一模式。本书是我此前的书《西方将主宰多久》的姊妹篇。在我写那本书时，编辑和我决定将其支撑材料放到一个网站上，而不是再出一本纸质书，但自那时起，情形就越来越明显，有些人希望得到这些材料的纸质版本，并希望材料更为广泛并得到修订。

我写《文明的度量》主要有两个目的。第一个目的是我想给批评《西方将主宰多久》的人提供些“炮弹”。他们需要这些“炮弹”来对我在该书中得出的结论进行系统的分析。尽管我当然希望我的论述能经得起这样的挑战，但如果我本人的分析能引发清晰的争论，最终得到更完善的社会发展指数，并对西方实力和财富的崛起做出更令人信服的解释，那么也是不错的结果。

第二个目的是想为比较史学变得更加清楚、更可量化做出贡献。由生物学家转为历史学家的彼得·图尔钦（Peter Turchin）曾指出：“科

学的历史是有说服力的。一门学科通常只有创造出数学理论，才是成熟的。”适合一切尺度，能回答所有从事比较研究的社会科学家们想问的所有问题的数值指数，是永远不可能存在的，但是能使比较史学成为这样一门成熟学科的最好办法之一，也许就是精心设计出多重指标，使每项指标都能解决一个特定问题。本书起首处，我要开宗明义地对我在提及“社会发展”这个概念时头脑所想的，做出正式定义。继而，我要对为做出这个定义所吸收的观点，以及最近50年来针对这些观点的异议，逐一综述。在第二章中，我试图从这些批评中提炼出社会发展指数所面临的主要挑战，然后解释我是如何努力应对这些挑战的。在本书的主体部分（第三章到第六章），我列举了我所取得的成果所依据的证据。这些成果主要是在能量获取、社会组织、战争能力和信息技术方面取得的，这是社会发展指数的4大特性。在最后一章，我思考了对于社会科学中的其他争论，社会发展指数有可能做出贡献的一些方式。

什么是社会发展

我所说的“社会发展”，指的是“社会达成目标的能力”。我要给这一特性贴上“社会发展”的标签，是因为它与发展经济学的核心观点有很多共同之处。历史学家彭慕兰（Kenneth Pomeranz）曾提出，将这一概念称为“社会力量”也许更好，但不能令我信服，尤其是因为这个概念与以前曾极具影响力的“社会力量”——特别是社会学家迈克尔·曼（Michael Mann）所定义的那种——极其不同，这个用词很可能会造成不必要的混淆。

社会发展是一个重要概念，因为西方（这是又一个需要定义的重要概念，在第二章有详细的说明）在过去200年来主宰世界的主要原因是：（1）其社会发展达到了比世界其他任何地方都高的水平；（2）这种水平之高，使得西方能够在全球施展其实力。

也许我们应当将“社会达成目标的能力”称作社会发展的最小定义。这样很便利，但却不够精确，而且像所有最小定义一样，其外延抽象化程度很高，使之很难用于操作（也就是说，我们没有必要立刻将这样含糊的一种说法付诸使用）。

因此，社会科学家们通常会在最小定义后续之以一个“理想型”定义。这种理想型定义“旨在汇集众多最大属性，也就是囊括所有（非特有的）有助于以其最纯粹、最理想（或许也是其最极端）的形式定义该概念的特性”。

用更正式的说法来说：“社会发展就是技术、生计、组织和文化方面的总体成就，人们通过这些成就来吃饭、穿衣、居住和繁殖自身，解释周围世界，在群体内解决纠纷，扩张自己的势力并压缩其他群体的势力，保卫自己、抵御其他群体扩张势力的企图。”

从原则上讲，社会发展是我们可以衡量并通过时间和空间进行比较的事物。如果西方的社会发展水平自远古时代起就高于世界其他地方的话，那么“为什么西方主宰世界”的答案一定深藏在历史中，正如西方生态和环境优越论的持有者们所主张的那样。然而，如果西方的社会发展是在公元前1000年后开始遥遥领先于其他地区的，我们也许就要断定：主张希腊、罗马重要性的人是正确的。但是如果事实证明西方的社会发展超越其他文明是发生在现代的事情，那我们就不得不承认那些旧的理论都是错误的，必须另寻解释了。

我想强调的一点是，社会发展是对社会达成目标的能力的衡量，而不是对社会达成目标的能力的解释。社会发展只是向我们展示了我们需要解释的模式。

社会发展也不是衡量不同社会的价值的标尺。例如，21世纪的日本，遍布着配有空调、实现了计算机化的工厂和繁华的城市，到处是汽车和飞机、图书馆和博物馆，拥有高科技的医疗条件和受教育程度

较高的居民。当今的日本人对自然环境和知识环境的利用远比他们1 000年前的祖先要充分。1 000年前的日本，上述事物一概没有。因此，要说现代的日本社会发展水平比中世纪的日本要高，是完全成立的。然而这绝不意味着现代的日本人比平安时代的日本人更聪明、更富裕或者更幸运（更不用说“更幸福”了）。同样，社会发展得分高低与道德、环境及其他社会发展的成本毫无关系。社会发展是一种中性的分析范畴。

进化、分化、能量：怎么理解社会发展

学者们对类似于社会发展的观点一直饶有兴趣。对于这一历史观点，已有若干出色的评述，所以我不打算再做全面概括了。我将只关注看上去与本书中将构建的社会发展指数最为相关的观点，然后介绍一些对这些方法最为重要的批评。

最有助益的起点也许当属那位古怪的英国大学问家赫伯特·斯宾塞（Herbert Spencer）于1857年发表于《威斯敏斯特评论》（*Westminster Review*）上的《论进境之理》（*Progress: Its Laws and Cause*）了。像19世纪中叶的许多英国知识分子一样，斯宾塞也感到自己生活在一个先前难以想象的进步时代，并想对此做出解释。他认为：“从科学所能探及的最遥远的过去，到新奇事物层出不穷的昨天，进步最本质的成分，是从相同性质转化为不同性质。”他提议将事物从起初简单到变得越来越复杂的机制称为“进化”：

> 事物经过持续分化的过程，从简单发展到复杂，可以看出在我们所能推论的宇宙的最早变化中，在我们能够通过归纳而确定的最早变化中，是相似的。这种发展可以从地球在地质和气候上的演变中看出；可以从地球表面每个单个的有机体的演变中，以及有机体的种类增长中看出；可以从人类的进化中看出，无论是关注文明的

个体还是各种族整体；可以从社会的演变中看出，就其政治、宗教、经济组织的变化而言；可以从所有那些或具体或抽象的人类活动的产物的演变中看出，正是这些无穷无尽的人类活动的产物构成了我们日常生活的环境。

在接下去的40年内，斯宾塞将地质学、生物学、心理学、社会学、政治学和伦理学归拢成一个涵盖一切事物的单一的进化理论，来解释宇宙是怎样从简单而无差别变得复杂而千差万别的。在其三卷本《社会学原理》（*Principles of Sociology*）中，斯宾塞提出人类社会的发展经历了4个演变阶段，从简单阶段（没有领袖的游荡营居群），经过复合阶段（有政治领袖的稳定村落）和加倍复合阶段（有教会、国家、复杂的劳动分工和学术活动的群体），到三倍复合阶段（像罗马，当然还有维多利亚时代的英国那样伟大的文明）。

斯宾塞的观点赢得了大量读者，自20世纪50年代起，很多人都认可了他的理论，并以此来构筑自己的思维，我将使用“社会进化论”这个术语，作为我本章将要讨论的所有理论流派的广泛的标签。我也将把“社会进化”（英国喜用的术语）和“文化进化”（美国喜用的术语）视为同义词。

到1870年时，斯宾塞已然成了用英语写作的最具影响力的哲学家。当19世纪晚期日本和中国的知识分子们认为他们需要了解西方的成功经验时，斯宾塞的著作也是他们的首选之一。甚至在《物种起源》（*Origin of the Species*）的前五版中并未使用“进化”一词的达尔文，在1872年修订第六版时，也不得不向斯宾塞借用了这个词。

还有几位19世纪末的理论家（经常和斯宾塞一起被称为“古典进化论者”）也创造了他们自己版本的斯宾塞类型学。例如，爱德华·泰勒（Edward Tylor）在其《原始文化》（*Primitive Culture*）一书中论述了从原始到野蛮再到文明的变化；路易斯·亨利·摩尔根（Lewis Henry

Morgan）在他的《古代社会》（*Ancient Society*）中也使用了同样的术语；这本书对恩格斯的《家庭、私有制和国家的起源》（*Family Private Property and the State*）产生了极大的影响。

可供这些理论家们利用的考古资料少之又少，因此他们非常倚重于假设。他们假设19世纪定居于非洲、亚洲、大洋洲和南美洲的居民，其祖先就是在那里居住的，他们的理论阐释了这些已经处于三倍复合阶段（即文明的演变阶段）的民族，一定是自史前时代就在那里定居了。然而，甚至他们所依赖的极其有限的人种学信息都充满了问题。这些信息大多来自传教士和殖民地官员，他们所感兴趣的，通常都只是他们所接触的族群的一些非常特别的特征。结果，当20世纪初，第一代专业人类学家独立展开田野调查后，他们很快就发现，那些理论家们所假设的事实，很多完全是错误的。

进入20世纪最初10年时，一股强烈的反冲发生了，在整个20世纪，斯宾塞关于进化和分化应当是历史研究的中心的观点，仅是昙花一现。其批评者中最重要的当属弗朗茨·博厄斯（Franz Boas，一位移居美国的德国学者）和布罗尼斯拉夫·马林诺夫斯基（Bronislaw Malinowski，一位移居英国的波兰学者）。到20世纪20年代时，他们已说服了很多人类学家，使他们相信田野调查的内容包括大量独立的“文化”，每种文化都必须理解为一个独特的、无缝的凝聚系统。

功能主义——认为观点、体系和价值会在所有这些独立的文化中达到平衡的理论——变得越来越流行，经常会打动人类学家们，使他们认为这是比那些古典进化论者们跳跃式的推测好得多的建设自然科学的社会学的基础。采用功能主义方法的代价之一，无疑是使得对随时间推移而发生的变化进行跨文化的比较和解释变得困难多了，但社会学家们通常愿意付出这个代价，而作为科学思维的一项组织原则的斯宾塞的进化论，很快崩溃了。

20世纪20年代，马克思主义者仍执着于用进化论来论事，但在自由民主国家（以及在法西斯国家，尽管相当困难），大多数社会学家和人类学家都认为，将人类群体按照从简单到三倍复合或从野蛮到文明的序列排列，就相当于编造既虚假又无意义的故事。

20世纪30年代也许是博厄斯单一主义的高潮期，但钟摆已开始向回摆动了。考古学家V·戈登·柴尔德（V. Gordon Childe），又一位移民的学者（这回是从澳大利亚移民到了英国），其学术生涯便极好地证明了这一点。在两次世界大战之间的年代，地层挖掘（即在一个考古遗址将沉积物分层，排列出可相互对照以确定时期的序列）成为考古学的规范，积累得足够多的资料使得广泛的综合分析成为可能。

柴尔德的第一部真正成功的著作《欧洲文明的开端》（*The Dawn of European Civilization*），简直是那个时代的典型，该书专注于某一特定地区，而不是像斯宾塞那样以全球为单位进行思考；以迁移和传播而不是进化和分化来解释文化的变化。然而到了20世纪30年代，柴尔德像自由民主国家的很多社会学家一样，转向了马克思主义，开始探讨一些非常不同的问题。在《人类创造自己》（*Man Makes Himself*）和《历史上发生了什么》（*What Happened in History*）两本书中，他意识到考古学日益扩大的资料库已经理所当然、不容置疑地展现出，在世界上的不同地区，农村和城市是独立发展的。到1951年时，他甚至打算写一本叫作“社会进化”（*Social Evolution*）的书。

就在同一时期，美国的许多社会学家也回到了进化论的框架之内。其中一些人像柴尔德一样，倾向于马克思主义，例如，人类学家莱斯利·怀特（Leslie White）以笔名发表了一系列左翼政论文章。不过也有一些人强烈反对这种倾向，例如，经济学家华尔特·罗斯托（Walt Rostow）给他的经典著作《经济成长阶段》（*The Stages of Economic Growth*）起了个副书名“非共产主义宣言”（*A Non-Communist*

Manifesto）。然而无论政治动机如何，美国人多倾向于像斯宾塞那样强调分化，而不是选择柴尔德的更人本主义的进化论。

这些思想者中影响最大的也许当属社会学家塔尔科特·帕森斯（Talcott Parsons）了。在一系列研究论文中，帕森斯不仅提出了一种社会阶段的新类型学[原始阶段、中间阶段（又可再细分为古代时期和晚近时期）和现代阶段]，而且提出了用以解释从原始到现代的发展的一种复杂的框架体系。帕森斯认为社会进化是由累进的6种“进化的共性”构成的，每种共性包括“各种结构形成的一种综合体，其相关发展过程会使特定种类的生命系统的长期适应能力获得极大发展，以致只有发展出这种综合体的系统，才能获得更高水平的一般适应能力”。首先产生的是社会阶层和文化正统性（即社会内的等级制度和差别，以及社会之间的群体认同性和差别），其次产生官僚机构和市场，最后是普遍性的规范准则（特别是在法律和宗教方面）和民主。

帕森斯的思想在于意图将从人类进化到20世纪的资本主义在内的所有一切，都囊括进一个单一的框架内，比柴尔德更雄心勃勃，但也因为其循环性，即认为分化既是进化的原因又是结果，而受到广泛批评。因此，尽管有一些社会学家认为帕森斯的理论所带来的冲击力是有趣的，但仍然转向其他方向去寻找对社会进化的解释。帕森斯之后，这些年来作品得到最广泛传诵的进化论者，似乎当属人类学家莱斯利·怀特了。他强调能量获取是推动进化的动力。像其他进化论者一样，怀特也将历史按阶段划分（他的划分法是：原始社会、文明社会和复合社会），然而与他的大部分前辈不同的是，他提出，“当每年人均利用的能量增长了，或者技术手段使能量发挥作用的效率增长了，或者上述两个因素同时增长了，文明就进步了。”怀特的结论是，历史可以概括为一个公式：C= E×T，即文明（Culture）=能量（Energy）×技术（Technology）。当人们从事农业生产后，便从原始社会进步到文明社会；实现工业化后便从文明社会进步到复合社会。

这是对斯宾塞–帕森斯路线的重大背离，但是当怀特关注起能量利用提高的结果后，他的理念被认为更接近于社会进化论的正统了。他认为，从原始社会通过文明社会到达复合社会的最重要的结果，是分化的增长。正如他所解释的：

> 农业……极大地增加了粮食供给，反过来又导致了人口的增长。随着人力劳动在农业生产中的效率越来越高，社会中越来越多的人力可以从获取食物这一任务中分离出来，转向其他职业。于是社会中出现了很多职业群体，如泥瓦匠、金属工、玉匠、织工、文书、牧师等。其结果是加速了艺术、手工艺和科学（如天文学、数学等）的进步，因为这些行业已掌握在专家，而不是“万金油”手里了。随着生产的发展，社会分工进一步细化，增加了更多的职业群体，生产的目的变成了交换和销售（而不再像部落社会那样主要是为了使用），于是出现了交换的媒介：货币、商人、银行、典当行、放贷者、奴隶等。财富的积累和对有利之地的争夺引发了征服战争，促成了职业军人和统治阶级、奴隶制和农奴制的产生。于是，存在于人力发展阶段的农业，使人类的生活和文化发生了深刻的变化。

第二次世界大战后的二三十年，美国思考社会进化问题的学者通常聚拢在“新进化论”的标签下，以便与（主要流行于欧洲的）19世纪的古典进化论相区别。新进化论的讨论大多贯穿着两大观点：一种是回归于分化是进化的最重要的结果（按照帕森斯的观点，也是进化的原因）；另一种观点希望对进化进行量化，以便做出更明确的对比。

用数值尺度来为社会进化分级的主张，可以追溯到19世纪晚期古典进化论的鼎盛时期。在可靠的、跨文化的数据的基础上进行这样的分级，最早的尝试也许当属泽巴尔德·施泰因梅茨（Sebald Steinmetz）的长篇大论《社会类型的分级》（*Classification des Types Sociaux*），是一篇主要关注生存技术的文章。汉斯·涅波尔（Hans Nieboer）在其

经典的论文《作为产业制度的奴隶制》（*Slavery as an Industrial System*）中更详尽地阐述了这种主张，伦纳德·霍布豪斯（Leonard Hobhouse）及其合作者则进一步扩展了这一框架。

到第二次世界大战结束时，美国社会学家又发现大量的新证据，统计技术也越来越成熟，使得那些早期的尝试显得毫无立足之地。人类学家卡尔顿·库恩（Carleton Coon）在一本大众教科书中所发的一番简短议论中，散布了这样的观点：通过统计一个社会中专家的数量、贸易量、企业集团的数量，以及机构的复杂程度，建立更加完善的数值指标应当是可能的，然而第一个真正可用的指标却是拉乌尔·纳罗尔（Raoul Naroll）提出的。

纳罗尔是“人类关系地区档案”（Human Relations Area Files，简称HRAF）的一名研究人员。“人类关系地区档案”是耶鲁大学于1949年建立的一个雄心勃勃的项目，旨在为人类行为、社会和文化进行全球性比较创建一个数据库。纳罗尔从全球范围内随意挑选了30个工业化前的社会（既有当前的，也有历史上的），然后遍搜“人类关系地区档案”以探究它们的分化情况。

鉴于分化可能存在几乎无限的维度，纳罗尔为这一概念的运用设立了两个原则。首先，他认为必须将研究限制于部分特性，这些特性能以最小的数目涵盖斯宾塞关于“分化的大部分想法”；其次，选定的特性必须符合一些基本的准则。它们必须有文化自由性（不能有种族优越的偏见）、逻辑独立性（不能充斥着伪相关）、充分的文献资料和可靠性（专家们不能对事实有太大的分歧），以及便利性（如果数据太难获得，则评分系统是不切实际的）。

纳罗尔将目光落在3个特性上：社会中最大定居点的规模、其手工业生产的专业化程度，以及该社会子群的数量。在调查了各种各样的定义问题和方法问题后，他对这3个特点进行了量化，将结果转换为标准格式，生成了一种“社会发展指数”，以63分为最高分数。在他的分

数表上，以12分垫底的是火地岛的雅甘人（Yahgan），就是在1832年令来访的达尔文惊叹为“生存于比世界上其他任何地方的人都要低的进化状态”的那群人；高居榜首的是15世纪的阿兹特克人（Aztecs），为58分。

几年后，当时还在美国自然历史博物馆任职的罗伯特·卡内罗（Robert Carneiro）提出了另一种迥然不同的建立指数的方法。卡内罗像帕森斯一样，也对下面这个问题感兴趣：所有社会在从一级复杂/分化水平提高到另一级时，是否都必须具备“进化的共性”（卡内罗称之为“功能性先决条件”）？他从社会心理学家们那里借来了量表分析技术，继而寻找起“具备下列特点的特性：（1）这些特性的出现标志着复杂性程度比它们未出现时提高；（2）这些特性一旦形成，就很可能保留下来，即使不是无限期地保留下来，至少也会在相当长时间内保留下来”。

卡内罗挑选了8个这样的特性（石建筑、政治状况、金属矿石冶炼、社会分级、织布机纺织、酿酒、陶器、农业），根据其有/无而打分，而不是像纳罗尔那样赋予数值。随后，他挑选了9个南美洲社会，将它们排列在一张他称之为“量表”的表中（见表1.1）。

卡内罗称，量表使他不仅能给这9个社会的复杂性评分，从0分（又是19世纪的雅甘人）到9分（15世纪的印加人），也使他可以证明这8个特性全部都是功能性先决条件，即“x必然优先于y，也就是说，如果没有x的预先存在，y是不可能存在的”。在后来的一篇文章中，卡内罗根据史实，检验了他的关于古代近东和盎格鲁–撒克逊的英国出现上述特性的顺序的指数，声称自己的方法可以自夸为高于0.90的“复制系数”。

表1.1卡内罗的量表，显示了南美洲9个社会中8个选定的文化特性的有（+）与无（−）

石建筑	−	−	−	−	−	−	−	−	+
政治状况	−	−	−	−	−	−	−	+	+

金属矿石冶炼	—	—	—	—	—	—	+	+	+
社会分级	—	—	—	—	—	+	+	+	+
织布机纺织	—	—	—	—	+	+	+	+	+
酿酒	—	—	—	+	+	+	+	+	+
陶器	—	—	+	+	+	+	+	+	+
农业	—	+	+	+	+	+	+	+	+
	雅甘人	舍伦特人	桂库鲁人	图皮南巴人	希瓦罗人	库马纳人	安萨马人	奇布查人	印加人

以会导致结果不同的不同统计技巧而进行的实验和得到的指数，在此后10年内大量涌现。其中大多遵循纳罗尔和卡内罗的模式，将反映大量不同的人类活动的特性聚拢在一起，试图得出对整个社会的概括印象，但也有一些人选择仅仅关注某种能够更直接地反映分化情况的特殊史料，比如丧葬情况或定居模式。然而，尽管有种种不同，各式各样的数据指数大多产生了近似的结果：根据卡内罗的统计，当时结论比较相近的分析家达到了87%~94%。

到了20世纪70年代后期，新进化论变成了非常热门的研究项目，部分要归因于两本书对基于分化的理论进行了非常清楚的阐释，分别是埃尔曼·瑟维斯（Elman Service）的《原始社会组织》（*Primitive Social Organization*）和莫顿·弗里德（Morton Fried）的《政治社会的进化》（*The Evolution of Political Society*）。前者将社会划分为营居群、部落、酋邦和国家，（更多地受到马克思主义影响的）后者则将社会分类为平等主义阶段、等级阶段、阶层化阶段和国家阶段。这些分类法（尤其是瑟维斯的分类法）多多少少在整个社会科学的领域里替换了帕森斯和怀特的术语。

20世纪70年代也许是美国新进化论的高潮。然而，就像是20世纪头10年——当时古典进化论眼看着就要产生出一种崭新的大综合——的离奇再现，在社会科学的许多领域，钟摆突然远离一切近似于斯宾

塞的理论而去。经济史和政治学是鲜有的例外，也许是因为制度分析日益增长的影响力推动了接近于过去的阶段理论，以及定量进化论在苏联依然受到青睐。然而在西欧和美国，对于进化论所进行的社会学、人类学和考古学争论，也像20世纪头10年那样带着政治锋芒。关于派性偏见、招摇撞骗和劣质学术的指责，败坏了20世纪80年代和90年代许多支持和反对进化论的著述的形象。

一些人类学家和考古学家认为“从简单到复杂的元叙事是一种主导性的意识形态。这种意识形态组织了当今世界史前史的写作，使之有利于现代化精神和西方优越性”；另一些人则回应，评论家们应当“抛弃他们对‘他异性’和‘自反性’之类的执着，转向评价真正重要的客观问题，进行一些深刻的思考，在解答问题时采用严格的定量方法”。大学里的人类学系往往是斗争最激烈的地方，一般分化为文化派和进化派两翼，有的各自招录教员和研究生（如哈佛大学），有的甚至索性分裂成两个系（如斯坦福大学）。

大约自2000年起，向社会进化论的又一次回摆似乎开始了。在20世纪70年代新进化论的鼎盛时期，自封为达尔文主义者的考古学家们是其最猛烈的批评者。按照主要的达尔文主义者之一罗伯特·邓内尔（Robert Dunnell）的说法，“如果说进化的意思是指其在科学中的作用的话，那么文化进化论既不是科学，也不是理论，更不是进化论。就其本身而言，在信守科学方法的考古学中，将其作为一个解释框架，是不适当的。”

然而，社会进化论的最新一次崛起，在很大程度上却是由生物学和社会行为的共同进化的理论化推动的。贾雷德·戴蒙德（Jared Diamond）的《枪炮、病菌和钢铁：人类社会的命运》（*Guns Germs and Steel: The Fates of Human Societie*s）一书，无疑是最具影响力的贡献。该书非常得体地将生物学、考古学、人类学和历史学研究结合起

来，引人入胜地讲述了过去15 000年来植物、动物和人类社会的共同进化。

戴蒙德在其学术生涯开端时是生物学者，曾在加州大学洛杉矶分校的医学院任教多年。现在他是加州大学洛杉矶分校的地理系教授。除了曾短期在斯坦福大学做过访问学者外，他从未在任何大学的人类学系、考古系或历史系担任过教职，尽管现在他已经成为这些领域最畅销读物的作者。

考虑到20世纪90年代关于进化论的学术争论中的火药味，戴蒙德的书之所以成功，很大程度上是因为吸引了非学术界的读者，该书是已经在大学外销售了数百万册后，才在大学内产生影响的，这也许并非偶然。这对于新的社会进化论来说似乎很典型。尽管再没有人写出能像《枪炮、病菌和钢铁：人类社会的命运》那样畅销的书，但政治学、经济学、宗教哲学、心理学、考古学、人类学和历史学领域的学者们都在为争取更广大的读者而写作。这一趋势使得大多数关于新进化论的著述摒弃了狭隘的专家腔调，又回归了斯宾塞和达尔文时代的风气，那时候严肃的学术著作都希望能为非专业读者直接阅读。

尽管学术界争论不断，我们还是有很好的理由相信，21世纪前10年有可能出现生物进化论和社会进化论的新的综合，论述这种新综合的著述将同时针对大学内外的读者。我写作《西方将主宰多久》和《文明的度量》，主要目的之一也是希望为这种新综合做出贡献。我提出的“社会发展”概念，脱胎于可回溯至斯宾塞的社会演变思想，建立在可回溯至纳罗尔的指数建设传统之上，然而本书仍将努力对在20世纪曾频繁再现的上述思想做出严厉批判。

在下一节中，我概括了一些对社会进化论的最重要的反对意见。我主要关注于过去50年，尤其是20世纪80年代的批评声音。这些批评可谓是这种方法最急需解决的问题。在本章结束之处，我从这些争论中提炼出了建立社会发展指数所必须克服的最重要的难题。

批评与缺点：谁犯了错

分化

我从“分化”这个词说起，是因为自斯宾塞以来的大多数理论家，都将这个词视为社会的维度，当社会进化得更为复杂时，其“分化”就会增长。然而，实际上，尽管“分化”现在已被普遍认为是核心概念，但经历了曲折的历史。

在分化问题上，考古学家们面临的困难也许比所有其他领域的学者都要大，因为他们发现分化实在是太难以度量了。20世纪70年代，考古界的一些社会进化论者被吸引去研究墓葬，希望丧葬礼俗能够清楚地表明社会分化出了什么样的社会角色，然而到了20世纪80年代，批评家们展示的是，埋葬者们通过他们对死者的不同待遇所要表达的，实际上是对活人之间应保持的理想关系的看法，而不像信奉帕森斯理论的社会学家们所认为的是实际社会角色的表现。于是，尽管分化在正式定义时很被看重，但在考古学家们实际判断进化/复杂性时，分化起不上太大作用。例如，在纳罗尔1956年的社会发展指数中，只有一个特性（定居点规模）会被理所当然地视为代表着分化；而在卡内罗最终的特性列表上，只有1/6的维度与分化直接有关。

自20世纪80年代起，考古学家们普遍渐渐放弃了将分化作为分析工具，然而社会学家们却越来越倚重分化。他们指出，在某种程度上，或许我们认为的最复杂的社会——伟大的现代民族国家——实际上在分化程度上还不及一些现代化之前的古代国家。这些古代国家有着复杂的住宅区网络、秩序和等级。查尔斯·蒂利（Charles Tilly）提出，分化是同质的市民社会兴起的标志。

这个过程对于现代性来说也不算独特：在另一个广为人知的事例中，公元前5世纪的雅典这一同质的市民社会，合法的分化程度也远不

及此前古代时期的城邦。尽管公元前6世纪和前5世纪的国家能力和繁荣程度有巨大增长，雅典墓葬所表现出来的状态结构的复杂程度却显著下降了，这点众所周知。

蒂利由此得出的结论是：“我们没有理由认为分化本身是一个连贯的、普遍的、像法律规定一样的社会过程。”于是，自20世纪80年代起，“分化”一词便从社会学的讨论中消失了，甚至比从考古学讨论中消失得还要彻底。

复杂性

如果说“分化”太不一致，无法作为社会进化理论的基础的话，“复杂性”一词也必须随之丢弃，因为在大多数社会学家的正式定义中，这个词是完全依赖于“分化”的。然而，在过去20年间，相当多的社会学家都提出，如果我们在斯宾塞的分化论基础上，以从自然科学领域汲取的复杂性理论来替代社会科学观念的话，“复杂性”可以保留为一个中心概念。

许多版本的复杂性理论都认为，如果我们将组织视为复杂的适应系统的话，我们很快就能看到宏观的模式和结构，从作用者依照完全不同的观念或者根本不依照任何观念行动的微观行为中显现出来。斯宾塞也许会欣赏这样的观点：秩序和等级制度的出现和毁坏是自然过程——通常被称为自组织临界性（Self-Organized Criticality，简称SOC）——对137亿年前宇宙的形成和人类组织的形成同样有重大意义。相关的观点已为人类学、考古学、管理学、历史学、国际关系学和政治学所采用。

复杂性理论的支持者经常会吸收和利用新进化论者的分类法，特别是当他们想要描述现代化之前的人类社会时。然而，他们也往往认为新进化论不严密、陷于经验主义和缺乏清楚的解释。

进化

虽然复杂性理论的支持者们复兴了斯宾塞的进化观，将其作为一个能涵盖从地质学到法律程序等一切的概念，但是自20世纪70年代起，有一些社会学家逆向而行，拒绝再使用同样的标签，以生物有机体的改变和社会组织中发生的多样化类型的变化来描述达尔文学说的派生理论。

最尖刻的批评也许是来自社会学家安东尼·吉登斯（Anthony Giddens）的。他提出任何自称为进化论的理论，“都必须至少与生物进化论有一些假定的概念上的连续性……（并且）社会进化论必须具体说明一些不只是一系列涉及某种特定条件的变化的东西，即某种能成为变化原理的东西”。他认为社会进化论与生物进化论的共性极少，特别是因为社会进化论依赖于将达尔文的变化原理——适应——不断扩展，直至其变得“无可救药地难以名状”。

许多生物学家都同意这一观点。约翰·梅纳德·史密斯（John Maynard Smith）尤其直言不讳。他是一位将博弈论运用于生物进化论的先驱人物。他认为“进化理论的解释力在很大程度上依靠三大假设：突变是非适应性的，后天属性是不可遗传的，遗传是符合孟德尔法则的——遗传是原子性的，我们平等地从我们的双亲那里继承原子，或者说基因，而不从任何其他人那里继承。而在文化类比中，所有这三大假设都是不成立的”。

一些考古学家对这类担忧的回答是，将人工制品视为人类表型的延伸，关注自然选择在它们古往今来的独特的持续状态中是如何运作的。罗伯特·伦纳德（Robert Leonard）总结了自称“信奉达尔文学说的考古学家们”的观点，解释说，“对社会进化论者来说，适应就是在环境中有功能的任何行为。对进化论者来说，适应是在很长时间内被自然选择所改变，以使其服务于某种重要的进化功能的一种表型特征。”信奉达尔文学说的考古学家们对于社会进化论往往比支持复杂性

理论的学者们更为苛刻。他们普遍认为社会进化论在关于选择的单位方面呈现无可救药的混乱，在论及适应时更是如此。

进步

今天已经很少有社会学家将“进步”这个词用作社会进化或分化的同义词了。但这是斯宾塞的核心概念之一，因此需要做一短评。

斯宾塞一定会认为诸如帕森斯的“进化的共性”和卡内罗的“职能先决条件”等社会学概念，与他的“进步”概念所见略同，无论后韦伯学派的社会学家们如何费力地试图将事实与价值分开，虽然许多社会进化论的批评者似乎也赞同。例如，考古学家迈克尔·尚克斯（Michael Shanks）和克里斯托弗·蒂利（Christopher Tilley）提出，关于进化、分化及相关概念的讨论，“很容易滑入自认正当或主张西方优先等思想方式”。如果他们的意见是正确的，那么关于进步的隐含假定就将不可避免地被纳入所有关于社会进化的讨论中。

阶段理论

实际上所有古典进化论者和新进化论者，从提出了从简单社会到三倍复合社会的类型学的斯宾塞，到提出了营居群、部落、酋邦和国家的交替论的瑟维斯，都创造了社会进化的阶段理论。这些理论有很多优点，尤其是它们具有能够预测无法直接观察到的变量的可能性。比如说，假如所有营居群都的确是以小股、流动的群落生活，人口密度很低，只有最低限度的技术，等级观念很弱，性别秩序肤浅，那么对于对一个社会只略知一二，比如只了解其生存基础和聚落形态的考古学家们来说，也许就能设想出该社会一些无事实证明的特征，比如其法律或其亲属关系。

根据HRAF的数据，考古学家查尔斯·麦克奈特（Charles McNett）声称推测准确率可达50%，而在20世纪70年代，很多史前史学者都在

辛勤地工作以阐明各阶段的考古关联，并在它们当中划分出具体的社会。然而，正如经常发生的情况一样，这种研究本身就会产生问题。个案研究发现一些社会并没有按照阶段理论所宣称的它们应该运行的方式运行；而对HRAF数据所进行的要素分析，又不能证明变量间存在明确的关联，因为不同的轮转会产生极其不同的负载。

20世纪80年代所进行的更仔细的跨文化调查表明，统计问题反映了真正混乱的现实。对新大陆各社会的调查发现“每个检查过的属性……都有相当大的可变性。这些差异是连续性的，而非孤立个别的，没有显而易见的社会模式或子类型。而且，不同的组织特征之间也发现了不同强度的关系”。

更糟糕的是，由于阶段之间清晰的分界线在现实世界中非常模糊，很难知道经验数据什么时候会证伪任何具体的阶段理论。在一个事例中，同一会议文集的撰稿者们，就人口密度和定居地规模是否与政治制度有确定的关联，得出了截然相反的结论。

一些考古学家为澄清问题，试图将瑟维斯提出的4个原始阶段划分出子类型，或者主张酋邦和国家代表着交替的发展道路，而不是连续的阶段。瑟维斯本人对混乱的回答是提出了一个更简单的“大分水岭”，在分水岭之前的“原始社会被划分为亲族群体，它们之间的关系是平等主义的”，在分水岭之后，“它们中的一些出现了等级制，由一股中央权威力量控制和指导——这股中央权威力量会创制成政府”。然而，大多数考古学家却持相反意见，越发认为阶段只是速记描述，或者为在不断变化的现实上进行启发而叠加的理想类型。

社会

伴随着对理论家们将社会聚拢在一起的阶段连贯性的挑战的，还有对“社会”本身的连贯性的挑战。

社会学家们很早就坚持认为“社会”是通过实践而非一元化制度组成的群体。人们也许会根据种族、政治、宗教、文化或其他条件来规定自己的社会。通常一个人会同时属于好几个社会，须根据具体情况以在它们当中做出选择（或者被选择）。迈克尔·曼将社会称为“同盟的、重叠的、交叉的网络”，吉登斯则谈及“像浅浮雕一样从它们嵌入的一系列其他系统性关系的背景中‘凸显’出来的社会系统。它们之所以凸显，是因为确定的结构原理足以创造出一种跨越时空的可具体指明的总体的‘制度的聚类’”。

人类学家也有这样的担忧。阿希尔·古塔（Akhil Gupta）和詹姆斯·弗格森（James Ferguson）为批判他们称之为“陈腐地‘在某某’模式”，在20世纪大部分时间里主宰了人种学的思想，提出：

> 无论地方和文化存在什么样的关系，都必须作为人类学研究的问题，而不能被当作一个人为讨论的出发点而做的让步；文化的领域化（比如成为种族或民族的文化）必须理解为正在进行的历史和政治进程的复杂且因情况而异的结果。正是这些进程，而不是先给定的文化领域实体，需要进行人类学研究。

社会学家们分析的“社会”，通常与人类学家们研究的“文化”殊为不同，似乎与考古学家们通常称为“文化”的人工制品的群集也不大相同。（考古学家们对“文化”的经典定义是：“在有限的地理范围内以聚集物的形式持续而反复地出现的明确而广泛的人工制品种类的多元集合。”）

纳罗尔意识到了这一问题，他创造了一个新术语“文化单位”作为回应。他将“文化单位”划分为4个类型，因两种地质年代计算尺度而有所不同，但这种复杂的想法没有得到很多人的支持。如果分析单位当真这样不稳定，那么社会进化论的基本内容，也就是长期大规模的比较，似乎注定要走向失败。

量化

量化对于社会进化研究的大多数途径都极为重要。半个世纪前，纳罗尔和卡内罗就已经为如何将名词转化为区间数据这一基本问题绞尽了脑汁。然而，到了20世纪70年代，将独特的人类状况或历史情境转换为可供计算的序列数据这一愿望本身就受到了挑战。正如尚克斯和蒂利所看到的，将历史“数学化”是进化论者们把西方主宰合法化这一不可告人的目的的一部分。他们指出，数学化背后错误的假设，就是当我们量化时，“我们重新发现了我们本来的数学上的自我，并且在对直接性和事实性的痴迷中发现了当前存在的必然性；这变成了客观必然”。

社会学家马克·格兰诺维特（Mark Granovetter）在其一篇经典的文章中提出，社会科学家们被拽向了两个相反的方向。一个倾向于将社会科学的主题“过度社会化”，在所有问题中都嵌入了太多的背景，给相互矛盾的意义建设和颠覆留下了太大的余地，以致任何解决方案都变成了不可能；另一个倾向是“社会化不足”，从背景中曲解细节，强加意义，因此得到的只能是肤浅的答案。

在抽象性和直接性之间寻找最好的平衡点就成为挑战。不同学科往往青睐于天平上不同的点，人类学和历史学在过度社会化方向上走得最远，而经济学和心理学在社会化不足方向上走得最远。如果（本章前面所引用的）彼得·图尔钦的说法“一门学科通常只有创造出数学理论，才是成熟的”是正确的，那么社会进化论就需要更多（也更精致）的量化；如果尚克斯和蒂利以及那些与他们持同样见解的人是正确的，那么数学化和社会进化论就不过是社会化不足的极端版本。

我们需要解释什么

这是一组可怕的批评攻势。如果这被证明是正确的，那么斯宾塞学派的社会进化论传统——以及通过一种社会发展指数来解答为什么西方主宰世界这个问题的所有希望——似乎都存在致命缺陷。

准备与批评者辩论以捍卫进化论的人为数并不少。不过，在本书中，我打算改弦易辙。过去半个世纪来出现的很多批评意见是很有道理的，值得认真对待，然而这并不意味着20世纪八九十年代出现的完全摒弃社会进化论的倾向也是正确的。在第二章中，我试图证明，如果我们认真对待批评，是有可能建立一种更集中、更具说服力的指数的。这个指数将能够避免新进化论的许多缺点，真正为我们提供一种工具，让我们看到如果我们想要知道为什么（目前）西方主宰世界，我们需要解释什么。

第二章

跨越时空的度量：社会发展指数

在第一章中，我提出，对于长达两个世纪的关于为什么西方主宰世界这一问题的争论，最好的解决办法是建立一个社会发展指数，因为这将使我们能够将西方社会的发展与非西方社会的发展进行长期比较。我们只有确定了需要解释的历史的轮廓，才能对为什么西方主宰世界做出更好的解释。

继而我查看了自19世纪50年代以来的社会进化论研究，以及自20世纪70年代以来对其最新版本——新进化论——的批评。在本章中，我将描绘一种社会发展指数。这种指数回应了对进化论的最严厉的批评，又没有忽视其中心目标——能够对社会发展进行跨越时空的度量和比较。

我的方法依赖于9个核心假设。首先，我会对它们进行阐述；其次，我将继续解释社会发展指数将如何运作。在本章结束时，我将解释为什么我认为这个指数完善了20世纪新进化论的指数。

定量、东西方、时间：什么支撑着社会发展指数

我做了9个基本假设，每个假设都需要开展不同程度的讨论。

定量分析

社会发展除非能够量化，否则就是一个无用的概念。已经有好几代历史学家对定量和定性的优缺点进行了争论，我不打算对这些越来越没有结果的争论老调重弹。

我并不想当然地认为定量分析就一定比定性分析更客观，无论我们是计算还是描述，都总会出现本能的判断和潜在的武断的识别。第三章至第六章将详细介绍一些在这方面，我所做过的最重要的判断和识别。

也就是说，定量分析的办法应当比定性分析更清晰，因为定量迫使分析者专注于这些判定，并且要系统地阐明做出某个选择而不是另一选择的理由。如果我们不对社会发展进行定量分析，争论就将继续陷于打概念仗的泥沼之中。所以我们的目标必须是为社会发展建立一个数字指标，使得对世界上的不同地区和不同历史时期进行直接的比较变得可行。

简约性

据说爱因斯坦说过："在科学上，应当使事情尽可能地简单，直到不可能更简单了。"尽管没有人曾试图寻找这句引言的原始出处，但我认为，必须以此作为研究社会发展的目标，但是并非所有人类学者都同意这样的看法。（就此而言，也并非所有自称社会科学家的学者都同意。）

学者们通常会认为做学问的目的应当是为我们对世界的理解增加复杂性。无疑有许多问题——尤其在文化研究方面——需要能使我们的认知更加复杂、更加玄妙精微的方法，为此甚至可以付出使之晦涩难懂的代价，但在讨论为什么西方主宰世界这个话题时，一般来说，主要是因为复杂性太重了，使得中心议题在大量细节中被模糊了，分析往往会陷入"只见树木，不见森林"的典型误区。

特性

设定像"社会发展"这样一个宽泛的概念，要求我们先将其分解为更小的、可直接度量的单位。遵照HDI模式，我试图在为社会发展的形式定义时涵盖全部标准的最小数量的具体特性。任何特性列表都不可能是完美的，但我们追求的目标是挑选出最佳集合。也就是说，如果我们在这种集合中增加更多的特性，就将经不起爱因斯坦的简单性的考验，因为那将使事情走向毫无必要的复杂；而如果我们减少特

性，那么这份列表就将无法包含定义的全部要素，并将使事情过分简单化。

第一个HDI是在1990年由经济学家马赫布卜·哈克（Mahbub Haq）设计的，旨在将发展经济学家们的注意力从国民收入核算上转移到人类福祉上。在和阿马蒂亚·森（Amartya Sen）及一个联合国经济学家团队的合作下，哈克巧妙地设计出了HDI，为联合国发展署的官员们提供了一种单一分数，使他们能够了解各国在促使其居民发挥其内在潜力方面做得怎么样。

HDI利用3个特性：出生时的预期寿命、知识和教育（以成人识字率占2/3的分数，以大中小学入学率占另外1/3的分数），以及生活水准[以美元计算的购买力平价换算的人均国内生产总值（GDP）]。联合国人类发展计划署经常改变其计算方法，2011年变动尤为明显，其进行了大改变，为生成分数提供了一个便利的计算器。

HDI被批评得体无完肤，从对特性的选择、基础数据的差错、对教育和收入的看重，到对环保和道德的忽视，都是如此，然而它仍然被证明极其有用，并得到了广泛运用。

人类发展与我在第一章中定义的社会发展并不相同，但是确定一个易处理的小数量的可定量分析的特性这一基本原则却是相通的。当然，HDI与我的社会发展指数也存在较大的差异，最明显的是，每个发表的HDI都是对一时的快照，不考虑历史变化。HDI可以通过比较每个年度报告中单一国家的分数，在一定程度上估量一段时间的变化，但由于其可能得到的最大分值总是1.0，HDI更适合于在图表中标示一个国家在某一时间点在世界上的相对位置，而不大适合于度量发展水平历经时间长河的变化。

总之，尽管HDI和我的社会发展指数目标大为不同，但其基本原则——小数量的可定量分析的特性能够作为较为宽泛的概念的代表物

——却是个极好的起点。

有用特性的标准

在社会科学中，就如何挑选好特性有很多讨论，大多数说法都关注6项标准：

（1）特性必须是相关的。也就是说，必须告诉我们一些与我在第一章中定义的社会发展有关的情况。

（2）特性必须是独立于文化之外的。例如，我们也许会认为文学和艺术的质量是衡量社会发展的有用的尺度，但是众所周知，对这些问题的判断是受到文化制约的。

（3）特性必须是相互独立的。例如，如果我们以一国人口数量和一国财富数量为特性，就不能再以人均财富量为第三个特性了，因为这是前两个特性的产物。

（4）特性必须有充分的文件证明。当我们回溯几千年前的历史时，这的确是个问题。特别是对遥远的过去，我们的确可能对一些可能有用的特性知之不多。

（5）特性必须是可靠的。也就是说专家们对证据的判断多少是一致的。

（6）特性必须是便利的。这也许是最不重要的原则，但取得证据越难，计算结果的时间越长，这个特性的用处越小。

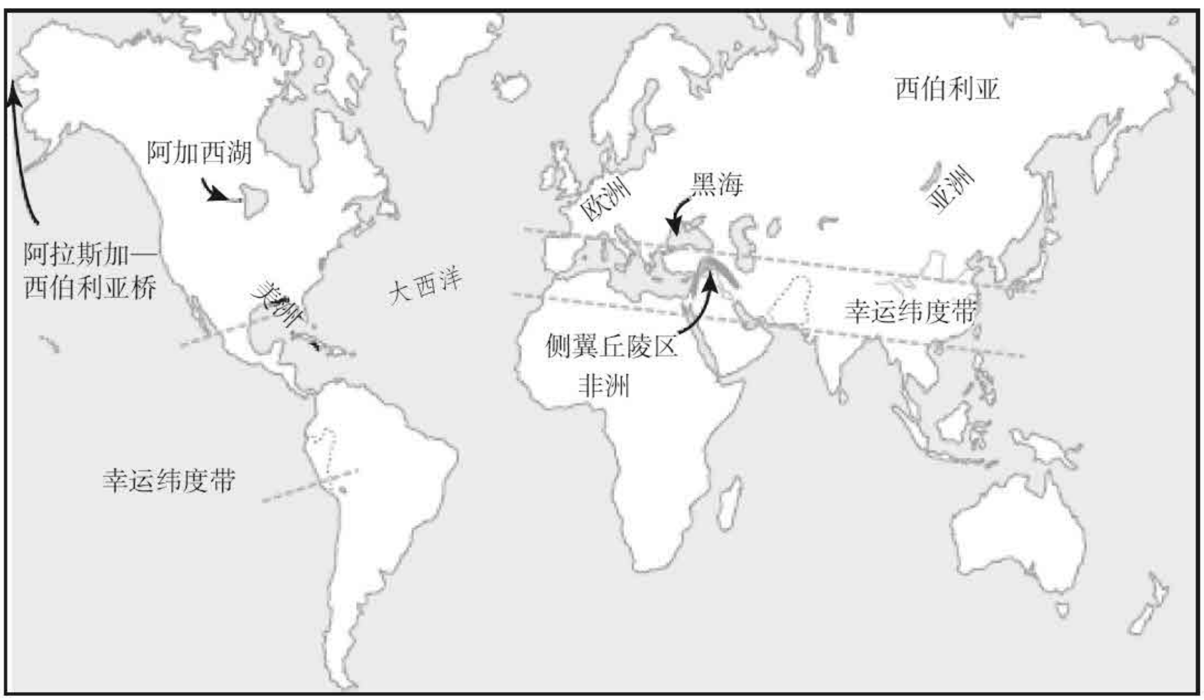
关注东方和西方，而不是整个世界

对社会发展进行真正的全球考察，尽可能详细地评估世界所有地区的情况，当然是非常受欢迎的。然而，那需要极其大量的工作，对

于解释为什么西方主宰世界这个问题着实是个钝器，它增添了不必要的复杂性，经不起简单的检验。

为什么西方主宰世界这一争论的核心问题，是西方的社会发展水平在遥远的古代就比世界其余地区的高，还是西方仅是在近代异军突起。为回答这个问题，我们没必要同样详细地调查世界上所有地区的社会发展状况。出于贾雷德·戴蒙德的《枪炮、病菌和钢铁：人类社会的命运》和《西方将主宰多久》第二章中讨论的原因，在最近一次冰期的末期，大约公元前13700年时，“幸运纬度带”（大致为北纬20°~35°的旧世界和南纬15°至北纬20°的新世界，见图2.1）的地区中的一小部分社会的社会发展开始加速，超越了世界上所有其他地区。

图2.1 幸运纬度带



资料来源：迈克尔·安吉（Michele Angel）制作

在过去数百年间，世界上仅有的似乎比较可能与西方匹敌的地区，是那些从新世界的核心地带发展起来的地区，以及南亚和东亚。这些地区在最后一个冰期的末期时，成为潜在的可用于家居的植物和

动物聚集最稠密的地区。实际上，自最后一个冰期末期起，唯一在社会发展方面得分高于西方的地区是东亚。因而，遵照精简原则，我关注东西方的对比。

哪儿是东方，哪儿是西方

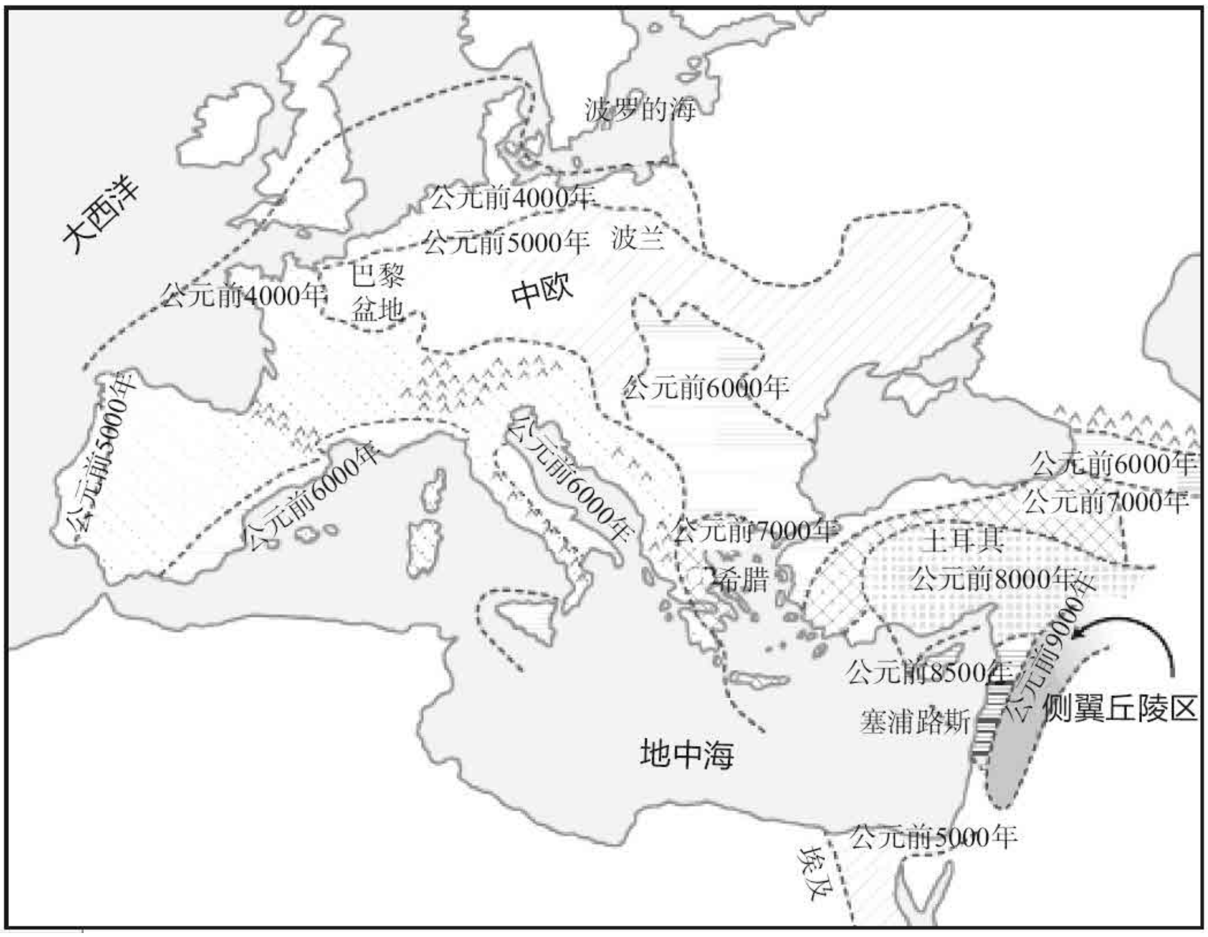
解释为什么西方主宰世界这个问题的最大困难之一，在于不同的学者倾向于以不同的方式定义“西方”这个词，使争论陷入了打概念仗的僵局。据历史学家诺曼·戴维斯（Norman Davies）计算，学术界共有不下12种不同的定义，相通之处仅仅是他所谓的“弹性地理”。戴维斯的结论是，西方“几乎可以被其定义者以任何他们认为合适的方式进行定义”，结果便是“西方文明本质上是个知识建构上的混合词，可用来进一步为其创造者的利益服务”。

戴维斯指出，问题在于历史学家们往往从他们喜欢与西方联系起来的一些价值着手，比如，民主、基督教、科学或自由，然后确定一批似乎享有这些价值的国家。然后再一本正经地将这些国家与一系列他们认为不享有这些价值的“非西方”国家对比，再就差异提出一种解释。他们的困境在于，反对的历史学家们也可以很简单地将另外一些价值定为西方的精髓，提取另外一些享有这些价值的国家进行不同的对比，自然会得出不同的结论，但能同样服务于自己的目的。

为了避免这种民族优越感，我做出了极其不同的假设。我不是建立一些我认为属于西方的价值，再沿时间向前回溯，而是一开始就着眼于故事开端，再向后展望。地区之间生活方式的根本不同，实际上只是在最后一个冰期结束后才开始发展起来的。当时一些群体开始栽培植物、驯化动物，而另一些群体仍在延续着采集食物的谋生手段。正如我在《西方将主宰多久》的第二章中所解释的，我对“西方”的定义是，人类在欧亚大陆最西部最早进行驯养生活的核心地带，即底格里斯河和幼发拉底河的源头地区，通过一系列殖民和竞争发展和扩张的社会。

在这一地区，驯养生活促使人口增长，同时推动了社会向前发展和人口向外扩张。到公元前4000年时，这个“西方”已经扩展至包括了欧洲大陆的大部，以及今埃及、今伊朗的西部边缘和中亚的一些绿洲（见图2.2）。在公元第一个千年，“西方”继续扩张，囊括了今天我们称之为“欧洲”的全部地区；在公元第二个千年，欧洲人又将“西方”扩展至美洲、大洋洲和非洲海岸。

图2.2 西方的早期扩张，公元前9000~前4000年



同样，当我提及“东方”时，指的是人类在欧亚大陆最东部最早进行驯养生活的核心地带，即黄河和长江之间的地区，也是通过一系列的殖民和竞争发展和扩张的那些社会。像在西方一样，驯养生活促使人口增长，同时推动了社会向前发展和人口向外扩张。到公元前2000年时，“东方”扩展至包括了我们今天称为“东南亚”的大部分地区。到

公元前1500年时，“东方”又囊括了今天的菲律宾和朝鲜半岛，在公元第一个千年中又兼并了日本。

这种将“东方”和“西方”定义为欧亚大陆最东端和最西端的驯养生活核心地带的社会的办法是常识。这样定义还有一大好处，就是使我们能够对长期的人类历史采取一贯的概念，避免困扰了很多关于为什么西方主宰世界的争论的意识形态极端化。

度量的时间间隔

社会发展指数的主要目标之一，是度量随时间推移而产生的变化，所以指数必须有历经时间长河的维度。我从最后一个冰期接近结束时，即公元前14000年开始打分，一直持续到公元2000年，这样不仅提供了一个便利的结束点，也使我们能再有一些时间看看此后的发展趋势如何。

遵照精简原则，在计算社会发展分数时，时间间隔应当小到足以显示变化的大致模式，但不能更小了。在史前时期，年代测定技术通常要允许较大幅度的误差，但社会变化的速度通常也非常缓慢。即使我们有足够好的证据来区别公元前12000年和公元前11900年，其差异也许仍然太小，无法度量。

因此我使用一种浮动的间隔。从公元前14000~前4000年，我以每1000年为单位来度量社会发展。从公元前4000~前2500年，证据的质量改善了，变化加速了，因此我以每500年为单位来度量社会发展。对于公元前2500~前1500年，我将间隔减为每250年为单位，最终从公元前1400~公元2000年，我以每100年为单位来度量。到了20世纪，资料的质量已使我们可以逐年探查变化。甚至如果我们愿意的话，至少自20世纪下半叶起，可以逐月查看变化，不过精确到这个程度对于回答为什么西方主宰世界这个问题并无多少助益，却要增加巨大工作量，违背了对有用特性的讨论中的第6项原则。

这个办法的缺点之一在于史前时代的变化不可避免地会被消除掉。的确，从长远来看，冰期结束后的头几千年，社会的发展变化比过去数百年缓慢得多，但在极少数的情况下，史前考古遗迹的年代也可以得到非常精确的确定（例如，法国阿尔卑斯山脚的湖畔村庄，通过树木年代学测出的时间，误差只在几年之间），很明显，这些漫长的波浪遮掩了很多较短的周期。目前，似乎还没有办法解决。

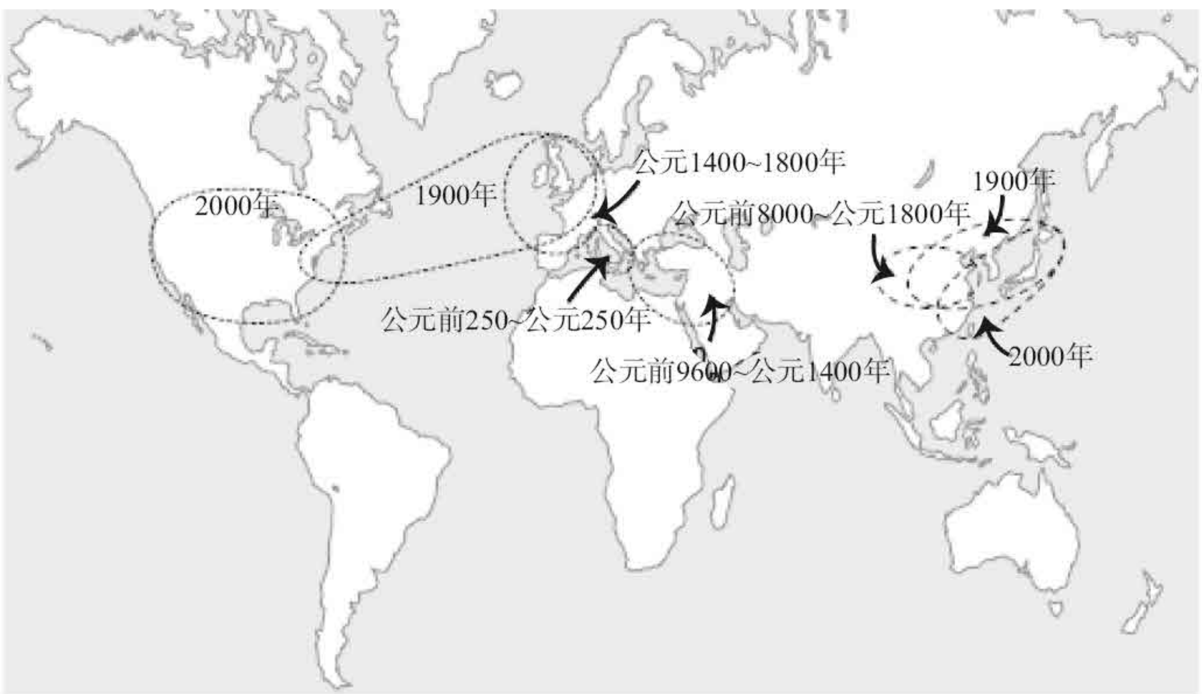
关注核心地带

历史学家彭慕兰在其里程碑式的著作《大分流》（*The Great Divergence*）中指出，一些历史学家热衷于提升欧洲的优越地位，于是经常犯一个低级的错误，即与将现代欧洲早期最发达的部分，如英国和荷兰，与整个中国对比，然后得出结论：欧洲在18世纪，甚至在17世纪，就要发达得多。将不相称的地区相比较，得出的是无意义的结果（这就是第一章中所提到的，纳罗尔提议用他本人创造的更抽象的“文化单位”，来取代考古学家们和人类学家们用于做比较的正式单位的原因）。因此，以适当的、可比较的时空单位来考察社会发展，对我们来说是至关重要的。

有一个解决方案，是把上述定义的整个东方地区和西方地区作为我们的分析单位，尽管这意味着西方的分数，比如说公元1900年的分数，将把工业化的英国、俄国的农奴，墨西哥的日工，以及澳大利亚的农场工人归在一起计算。继而我们不得不为整个西方地区计算一个平均的发展分数，然后再对东方重做一遍同样的工作，再对历史上每一个更早期的点重复同样的进程。这样做就太复杂了，因而是不切实际的，违背了对有用特性的讨论中的第6条原则；而且这样做也许根本是无意义的。当要解释为什么西方主宰世界时，最重要的信息通常会来自比较每个地区最发达的部分，也就是聚集着政治、经济、社会和文化相互作用最强烈的核心地带。社会发展指数需要衡量和比较这些核心地带内发生的变化。

随着时间推移，这些核心地带发生了极大的迁移和变化（见图2.3）。从公元前11000~约公元1400年，西方的核心地带在地理上非常稳定，一直安于地中海的东端，除了约公元前250~公元250年的500年间，罗马帝国向西扩张，纳入了意大利。否则，西方的核心地带就始终维持在今伊拉克、埃及和希腊形成的三角形内。自公元1400年起，这一核心地带不断地向北和向西迁移，先到达意大利北部，继而到达西班牙和法国，然后又扩张至英国、低地国家和德国。公元1900年，该核心地带跨越了大西洋。到2000年时，牢牢地在北美洲扎了根。

图2.3 东方和西方核心地带的迁移



资料来源：迈克尔·安吉制作

在东方，核心地带始终保持在最初的黄河–长江之间地带，一直到公元1850年。尽管大约公元前4000年后，重心北移至黄河流域的中原地区，公元500年后又回到了南方的长江谷地，公元1400年后再次逐渐北移。到公元1900年时，该核心地带纳入了日本，到公元2000年时又纳入了中国东南部。

至于东方和西方的核心地带的确切边界在哪里，无论在任何时间点，专家们都会有不同见解。在表2.1中，我大致标示了核心地带。

表2.1 公元前14000~公元2000年西方和东方核心地带

西方	
公元前14000~前5000年	侧翼丘陵区（亚洲西南部）
公元前4000~前3500年	美索不达米亚平原（亚洲西南部）
公元前3000年	埃及（非洲东北部）
公元前2500~前1300年	埃及（非洲东北部）、美索不达米亚平原（亚洲西南部）
公元前1200~前1000年	埃及（非洲东北部）
公元前900~前700年	叙利亚–美索不达米亚平原（亚洲西南部）
公元前600年	埃及（非洲东北部）、美索不达米亚平原（亚洲西南部）
公元前500年	波斯帝国（亚洲西南部）
公元前400年	波斯帝国–爱琴海（亚洲西南部－非洲东北部－欧洲东南部）
公元前300年	希腊化王国（亚洲西南部－非洲东北部－欧洲东南部）
公元前200年	地中海盆地（亚洲西南部－非洲东北部－欧洲东南部）
公元前100~公元200年	地中海中部（欧洲南部）
公元300~600年	地中海东部（亚洲西南部－非洲东北部－欧洲东南部）
公元700~800年	埃及（非洲东北部）、叙利亚－伊拉克（亚洲西南部）
公元900年	埃及（非洲东北部）、西班牙（欧洲西南部）
公元1000~1400年	地中海盆地（亚洲西南部－非洲北部－欧洲南部）
公元1500~1600年	大西洋沿岸地区（欧洲西部）
公元1700~1800年	法国、英国（西北欧洲）
公元1900年	德国、法国、英国、美国（欧洲北部，美洲北部）
公元2000年	美国（美洲北部）
东方	
公元前14000~前2250年	黄河－长江河谷（中国）
公元前2000~前600年	黄河河谷（中国）
公元前500~公元200年	黄河－长江河谷（中国）

公元300~900年	长江河谷（中国）
公元1000~1500年	黄河－长江河谷（中国）
公元1600年	黄河－长江河谷（中国）、日本
公元1700~1800年	黄河－长江河谷（中国）
公元1900年	日本
公元2000年	中国东部、日本

近似和讹误

理所当然的是，世上根本没有百分之百精确的指数，无论是从较严格的意义上来解释“精确”一词，使其意味着所有单个细节都绝对正确，还是从较宽泛的意义上解释，使其意味着所有专家都会做出同样的估计，即使他们无法证明这些估计是正确的。在所有历史知识中，我们可以完全确信的只有很少一部分，而专家们能一致同意的更少。因此，如果问我计算的社会发展分数是否正确，是没有意义的。它们当然是不正确的。唯一有意义的问题是，它们有多不正确？它们是否大错特错，证明了我错误地识别了社会发展历史的基本形状，意味着我对为什么西方主宰世界所做的解释存在致命的缺陷？抑或这些误差实际上都微不足道？

解决这些问题有两个主要办法。一个办法是假设我犯了系统错误，普遍地高估了西方的分数而低估了东方的分数（或者相反），然后再问：（1）我们需要将分数改变多少，以使历史显得非常不同，使得《西方将主宰多久》一书中所提出的论点不再适用？（2）这样的改变是否貌似合理？我将在第七章中回答这些问题。

另一个办法是假设错误不是系统性的，高估或低估东方和西方分数的错误都是以随意和不可预知的方式犯下的。解决这类问题的唯一办法是自我修复我个人打分所基于的证据。我将在第三章到第六章详细说明这种办法。

社会发展是可以计算的

第一项挑战是找到符合有用特性假设下所列6项标准的最小数目的特性。在试过几种组合之后，我选定了4个特性：（1）能量获取；（2）社会组织；（3）战争能力；（4）信息技术。

能量获取必须作为社会发展的所有可用的衡量尺度的基础。牛顿的热力学第二定律告诉我们，物质的复杂组合使得其如果没有自身所处环境的补充能量的输入，经过一段时间就会分解。人类如果不获取能量，也会（像植物和其他动物一样）死去。同样的，除非人类能从所处环境中获取能量，否则人类创造的社会也会瓦解。人类群体为了增强对他们所处的自然环境和精神环境的控制，把该做的事情做好，不得不增加其能量获取。

然而，单是能量获取不足以衡量对社会发展重要的一切事物。甚至我所知道的对文明最简化的定义——莱斯利·怀特的C= E×T，都认为衡量人们对他们获取的能量的使用方式，与衡量能量本身同等重要，这是理所当然的。然而，怀特对“技术”的分类太过宽泛、太难以量化，所以我进一步将其细分为3个特性。

社会组织是其中的第一个。这一概念不可避免地在相当大的程度上与斯宾塞的分化概念重合，但为了避开我在第一章中提到的关于定义和衡量尺度的无休无止的争论，我从经济学家那里学来了一招，以社会中最大永久定居地的人口规模，作为对社会组织的大致的间接度量。

这也许看上去像种奇怪的方式。今天世界上一些最大的城市简直是功能失调的噩梦，充满了犯罪、污秽和疾病。然而历史上的大多数大城市无疑也是如此。在公元前1世纪，罗马有100万居民。那里也有街头犯罪团伙，时常会导致政府停转；城市人口死亡率过高，以至于

每个月都不得不容许上千名乡下人移居罗马，以保持人口数量。然而尽管罗马有种种邪恶，保持其城市运转的组织，却是世界上任何其他早期社会所远远无法胜任的——正如管理像拉各斯（人口1 100万），或孟买（人口1 900万），更不用说东京（人口3 500万）这样的城市，要求有远高于罗马帝国的组织能力。

这就是社会科学家们经常将城市化作为评估社会组织的大致尺度的原因。这样做有好几种办法。我们可以计算一个社会中居住在特定规模的定居点（10 000人是较流行的截断点）的人口的比例，也可以将定居点分成不同的等级，计算每个社会有多少个等级。然而我选择的办法，是只分别计算东方和西方社会最大的永久定居点的人口数量。我选择这种办法，一是因为如果我们要将研究一直回溯到公元前14000年的话，我们必须使用一些证据，而这种办法看上去最适合于这类证据；二是因为我知道还没有哪项研究表明这种办法比任何复杂得多的办法更不好用。

信息技术是能量使用方面又一个不可或缺的要素。随着社会发展的推进，人们必然要处理和交流巨量的信息。任何社会如果没有文字和计算系统，都不可能发展得太远。为了发展得更远，要求有越来越成熟的媒体储存和传播信息，要求有专门的机构将读写和计算技能传授给越来越多的人。

战争能力也是社会发展至关重要的一部分。像植物和所有其他动物一样，人类若想生存，必须竞争和合作。鉴于人类（像蚂蚁和黑猩猩一样）已经发展为一种社会性物种，需要时常将其合作性的行动转为群体针对其他敌对群体的暴力竞争。在考古记录中，武器和堡垒占有非常突出的地位；当信息技术达到能够记录细节的阶段后，世界上大多数文明的文字资料中，都充斥着对战争和战役的描写。

这4个特性并非合起来就构成了过去16 000年社会发展的总体画面，也没有比联合国的特性——预期寿命、教育和收入——告诉我们

更多的关于我们想了解的人类发展情况。这些特征的功能非常有限：它们只是给我们提供一个关于社会发展的有用的概览，展示一下我们在回答为什么目前西方主宰世界这个问题时需要解释的模式。

并不是只有这4个特性能起到这样的作用。我也观察过若干其他可能的特性，包括最大的政治中心的人口规模、科学能力和对于技术能力的更广泛的衡量尺度，但是，没有一个能像能量获取、社会组织、信息技术和战争能力这样，更好地服务于有用特性假设下开列的原则。也就是说，特性中的确有相当大的冗余，说明与社会发展的核心概念关联密切的任何特性组合，都可能产生非常近似的一套分数。

250分还是0.01分：哪个社会更发达

对于任何指数来说，最大的挑战无疑都是决定如何给特性分配分数。为了使情况简单，我决定将1 000分定为指数的顶点——公元2000年前所可能得到的最高分数。然而，这个顶点与HDI中可能的最高分1.0所起的作用非常不同。在联合国的指数中，1.0代表着某种完美，意味着任何一个社会的得分都不可能高过1.0。而我的社会发展指数则相反，1 000分只是公元2000年所能得到的最高分。从指数的顶点到本书写作时，又过了十几年，西方发展的分数在继续上涨，已经超过了1 000分。假如东方和西方继续以20世纪的速度发展，到21世纪末时，两者都将达到5 000分；而如果增长率超过20世纪的速度——目前的情况似乎正是这样——两者在2100年时的得分还会更高。

设计了HDI的经济学家们创建了一个详尽的权重体系，将3个特性的分数标准化，然后计算出相加后总分的平均数。相反的是，当1956年纳罗尔发表了他原创的社会发展指数时，他给他的3个特性设置了相同的权重，他解释说，“因为没有明显的理由认为某一特性比另一特性有任何更高的重要性。”

不同权重的优缺点总是有讨论的余地，我将在第七章再回到这个问题上，但是纳罗尔的办法在这里似乎比联合国的要中肯。即使有很好的理由确定某项特性比另一项意义更重要，也没有根据假设这一重要性能贯穿我们所考察的整个16 000年，或者它们始终同样适用于东方和西方。

因此我把我的1 000分平均分配在4个特性上。这意味着在某个特性上能获得最高值的社会，将在其达到那个水平的阶段（在任何情况下，都将是公元2000年）获得250分，而其他社会因为达到的是较低值，相应地获得较低的分数。我将在第三章到第六章详细说明证据、定义问题和我如何计算每个特性的分数，但在此，我要简短地举一个具体例子，以说明打分系统的运行机制。我将以社会组织为例，通过最大定居点的规模来进行间接度量，因为这也许是最简单明了的特性了。

大多数地理学家都将公元2000年时的东京，归为自公元前14000~公元2000年我们所知的最大城市，该城有大约2 670万居民。于是，公元2000年的东京获得了划拨给社会组织这一特性的满分250分，即意味着在指数中，每106 800人得1分（2 670万人除以250分）。公元2000年，西方核心地带的最大城市是纽约，有1 670万人。按照每106 800人1分的算法，纽约的得分是156.37分。

公元1900年的资料没有这么完备，但所有历史学家都一致认为那时候的城市要小得多。在西方，1900年的伦敦大约有660万居民，得61.80分；而在东方，最大的城市还是东京，但那时只有175万人，得16.39分。

当我们回溯到公元1800年时，历史学家们不得不将若干不同种类的证据结合起来，包括食品供应和税收记录、城市覆盖的自然区域、这些区域内房屋的密度，以及逸闻故事等，不过大多数人的结论都

是，北京是当时世界上最大的城市，人口可能有110万，得10.30分；西方最大的城市还是伦敦，人口有大约86.1万人，得8.06分。

沿时间继续回溯，误差幅度将越大，但在公元1700年之前的1 000年间，世界上最大的城市显然在中国（日本的城市经常紧随其后）。公元800~1200年，首先是长安，其次是开封，最后是杭州，人口在100万上下（分数在9分左右）。与之相反的是，西方城市始终没有超过其一半的规模（西方最大的城市通常在欧洲南部和亚洲西南部的伊斯兰地区，而不是欧洲北部和西部的基督教地区）。再往前几个世纪，这种情况则要颠倒过来：公元前1世纪，有100万人口的罗马无疑是世界上最大的大都会，而中国长安的居民人数可能只有其一半。

当我们回退到史前时代时，证据无疑越发模糊了，数字也变得小得多了。然而，将系统的考古调查结果和较小区域内详细的发掘记录结合起来，我们对于城市规模仍能产生合理的感知。这在很大程度上是分解的艺术，不过就算是最广为人们接受的估量结果，仍可能需打10%的折扣，但误差幅度不可能更大了；而且由于我们对东方和西方的考古遗址采用同样的估算方法，大概的趋势应该还是相当可靠的。

由于在社会组织这个特性上，需要有106 800人才能得1分，略微超过1 000人将得到0.01分，这是值得纳入指数的最低分数了。能达到这个水平的西方最大的村落出现在大约公元前7500年，而东方最大的村落出现在公元前3500年左右。在这两个时间之前，西方和东方的得分都将为0（参见第四章）。

另外3个特性的分数，也以同样的方式计算：（1）确定在这个生活维度中能得到最高分的社会（就能量获取而言，在公元2000年的美国，平均每个居民每天消耗大约228 000千卡的能量）；（2）将满分250分分配给这个社会；（3）计算能得到1分所需要的表现（就能量获取而言，228 000除以250分，等于平均每人每天消耗912千卡得1分）；（4）估算历史上不同时间每个社会在这个特性上的分值；

（5）将这些估算出的分值除以分母，将它们换算成指标中的分数。在计算了自公元前14000~公元2000年整个时期中每个特性的得分后，我只需将4个特性的得分相加，就能得到一系列每个地区社会发展的分数。这就使得我们可以将世界上不同地区在历史不同时期的社会发展状况进行对比了。

反对者的声音

根据第一章回顾的关于新进化论的争论，可能存在4种主要的反对我的意见。我将针对每一种说上几句，解释一下为什么这几种反对意见都不是致命的。

1.将不同时代和地区的社会发展量化并进行对比，使人类失去了人性，因此我们不能这样做。

这是20世纪七八十年代人类学和社会学领域反对新进化论的浪潮中最具影响力的声调，类似的观点在历史学家中也赢得了许多追随者。然而，这却是至今提出的反对意见中也许最令人无所谓的，因为一旦我们认识到不同的问题需要我们以不同程度的抽象性来解决，其力量就基本消失了。

到20世纪80年代后期，许多领域的学者们都感觉到像新进化论、新古典经济学等高度抽象的科目，以及同源方法，都留下了太多未解之谜，于是他们非常明智地转变了观念，采取了似乎更有利于做好工作，以回答他们感兴趣的问题的思维方式。例如，许多社会学家不再将支持常态和结构化的分化和功用作为组织概念，甚至社会科学中最坚定地致力于量化和抽象化的人口学和经济学，也进行了自身的后现代转型。

当然，我发现了我自己的研究正是如此。在利用新进化论和20世纪80年代中期的比较框架，弄清了铁器时代希腊的社会变化后，这些方法的局限性就变得越来越明显了。古希腊社会一些最重要的特征，如激进的男性民主制与大规模奴隶制的结合，很不符合瑟维斯、帕森斯和弗里德的理论，他们都认为民主是与先进的现代国家相关联的。如果完全不解释希腊民主制的来源，无视古希腊人的独特成就，而在社会发展指数中给希腊打分，就会使事情更加困难。

但这并不意味着社会进化论和社会发展指数都是在浪费时间，这只是意味着回答这个特别的问题还需要其他工具。用一种更狭隘、更排他主义的方法去研究古希腊社会，会比进化论的方法产生更多成果，但这样也有局限性，尤其是不足以解释经济、军事和政治发展是如何驱动了公元前第一个千年内的变化的。为了弄清这些问题，也得到其他材料的指引，我重新转向更广泛的进化论工具和将希腊置于全球框架中的需要。

问为什么目前西方主宰世界，与问为什么有些希腊城邦给予所有成年男性公民投票权，是不同类型的问题。这是一个宏大的比较研究的问题，要求我们纵观数千年历史，横跨百万平方千米土地，汇聚几十亿人口。为实现这一目标，一个社会发展指数恰好是我们需要的工具。

2.将社会量化并进行比较是一种合理的做法，但我定义的社会发展（社会达成目标的能力），却不是应该度量的事物。

这种反对意见用不着多费口舌就能解决。持这种意见的批评者需要表明有其他事物可供度量和比较，对于解释为什么目前西方主宰世界，将比我定义的社会发展更加有益。我不知道是否存在这样的其他事物，所以我请批评者们来确定，并证明它们能产生更有益的结果。

3.我所定义的社会发展，也许是一种在历史长河中比较不同地区的有用方式，但我用来度量的特性（能量获取、社会组织、战争能力和信息技术），并不是最好的特性。

这种反对意见可能以3种形式出现:

（1）在我考察的4个特性（能量获取、社会组织、战争能力和信息技术）之外，我们还应增加更多的特性。不过尽管必然还有很多特性可供我们考察，但精简原则要求我们在能涵盖社会发展的全部意味的最小数量的特性之外，不要再增加更多的特性。批评者需要证明我的4个特性实际上未能涵盖社会发展的一个或多个重要方面，而涵盖这些方面将会产生足以与我的指数大相径庭的结果，因而值得付出额外的努力，增加工作的复杂性。

（2）我们应当采用不同的特性。同样，当然有其他变量可供我们度量，但我考察过的所有其他特性，都不大符合前述各种原则，普遍存在严重的经验主义问题，或文化依赖，或相互重叠。如前面所提到的，大多数其他特性在历史大部分时期中都多多少少显示出相当大的冗余，而对于它们的任何貌似合理的组合，往往都会产生大致相同的最终结果。

（3）我们关注的特性应当更少。考虑到4个特性当中存在的冗余，我们或许应减少几个特性，以增强简约性。很显然，方法也许是放弃社会组织、战争能力和信息技术，只专注于能量获取，因为社会组织、战争能力和信息技术都只是使用能量的途径。图2.4展示了单看能量的指标会是什么样的。图2.5则不同，展示了由全部指标产生的分数，但两者的区别并不大。在单有能量的图中，正如在完全的社会发展图中一样，自冰期晚期以来，西方仍在90%的时间里领先东方；东方仍然是在大约公元550~1750年超越了西方；大约公元100~1100年仍然有一个硬上限阻碍了发展（稍高于每人每天30 000千卡）；后工业

革命时代的得分仍高于较早的时代；公元2000年仍然由西方主宰世界。

图2.4 公元前14000~公元2000年东方和西方的能量获取情况（线性标尺）

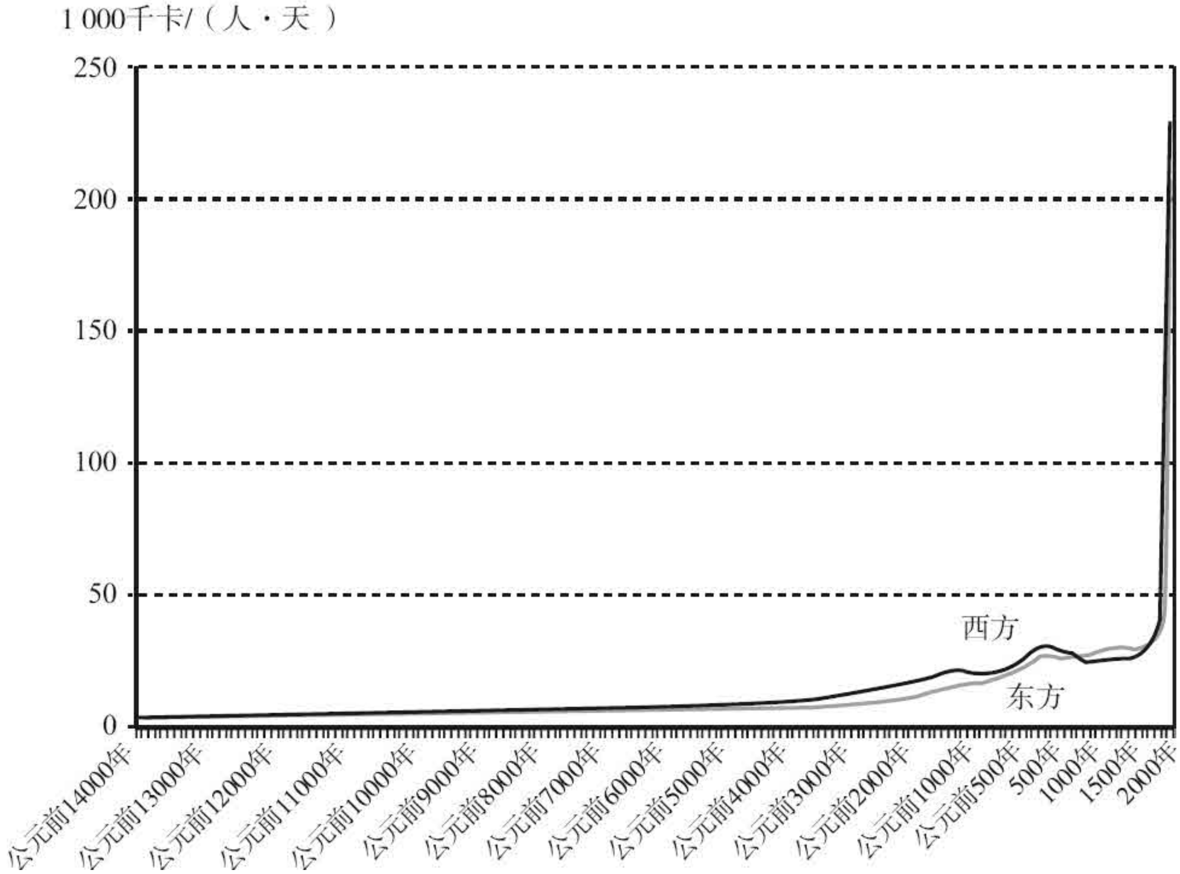
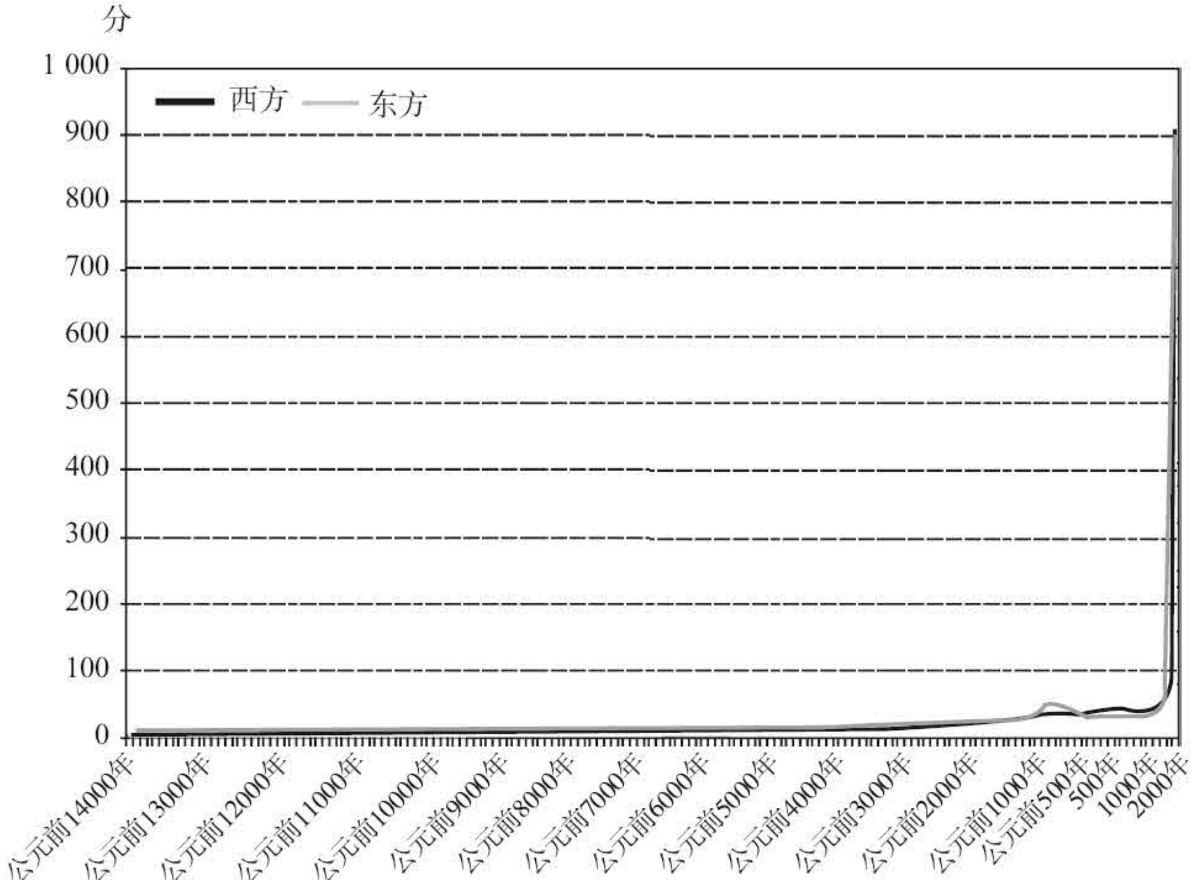


图2.5 公元前14000~公元2000年东方和西方社会发展得分（线性标尺）



单是关注能量，当然有更加简约的好处，但也有巨大的缺陷。我采用的4个特性并不完全多余，自大约公元1800年工业革命开始以来，能量获取与其他特性呈现非线性关系。能量获取的余裕的增长，导致了选定领域能量使用的极其巨大的增长。新技术的出现，20世纪城市的规模翻了4番，战争能力增强了50倍，信息技术跃进了80倍，而人均能量获取才翻了一番。如果只关注能量，就经不起爱因斯坦理论的检验，因为那样太简单了，扭曲了历史的原貌。

4.这4个特性是度量社会发展的好办法，但我犯了事实错误，得到了错误的度量结果。

正如在讨论近似和讹误时提到的，应对这个反对意见有两种主要方法。一种是假设我犯了系统性的经验主义错误，始终高估了西方的分数而低估了东方的分数（或者相反），然后问两个问题：（1）我们

需要把分数改变多少，才能使历史看上去非常不同，以致《西方将主宰多久》一书中提出的观点不再有效？（2）这样的改变是否似是而非？我将在第七章里论述，有充分的理由认为这种意见不能成立。

另一种辩驳这种意见的办法是，假设存在一贯但并不系统的经验主义错误，随意、出人意料且严重地夸张或低估了东方和/或西方的分数。避免这一危险的唯一办法，无疑就是贯彻第三章到第六章提出的意见，对照证据检查指标中的分数。

通吃：展示历史的轮廓

对于我创造的社会发展指数，我最想说的是，它反映了对社会进化论的批判，也同样反映了社会进化论者们的贡献。从斯宾塞最初的文章发表以来，社会进化论的批评者们就辩称，社会进化论者们试图解释一切，却往往最终什么也没有解释清楚。

批评者们有一点显然是正确的，就是没有能通吃的社会发展指数。纵观整个20世纪，指数创造者们想把各种各样的主题纳入自己的框架，却往往使得不具备可操作性，而解决这个问题的努力，经常又使情况雪上加霜。例如，卡内罗的解决办法是给他的指数增加越来越多的特性，结果，特性从1962年的8种，膨胀到1970年的618种，仍然还望不到明显的尽头。

本书中的指数并非试图解释一切，只专注一个问题：为什么欧亚大陆西端的社会，在19世纪主宰了世界，而其在北美的殖民地，又在20世纪取代了它们。

这样的专注有3点好处。它使得3种情况成为可能：（1）定义社会发展的核心概念时，头脑中有这一特定问题；（2）选择能直接证明核

心概念，同时又合情合理、易于操作的特性；（3）设计能在时间的推移中度量变化的指数。

这些好处也使得我避免了许多摧毁了新进化论者的指数的困难。其中最重要的也许是棘手的分化概念。这个概念是从斯宾塞那里传承下来的，但几乎不可能付诸实践。由斯宾塞学说的分化概念发展而来的某些见解，在任何关于社会发展或社会进化的有用的定义中，都应占有一席之地。而这种见解的确在我的指数中出现了，作为更广泛的社会组织特性的一部分，通过城市规模而进行间接度量。然而，这里提供的指数确保不会陷入麦圭尔（McGuire）在新进化论者对分化的研究中发现的那些陷阱。

同时，这个指数也避免了与任何特定的社会进化理论密切结合（相反的是，比如说，与卡内罗的量表分析倒是密切结合，而量表分析又是与直线发展的阶段理论明显相关联的）。这个指数也同样便于度量是否所有社会都的确是沿着卡内罗提出的路线发展的，或者自组织临界性是否始终符合长期、大规模的社会变迁。

本章介绍的方法对于解决分析单位问题，也有一定的作用。这个问题曾令20世纪的社会进化论者痛苦不堪。这里介绍的办法，也使得在东方和西方各确定一个核心地带，有了很大的灵活度，由社会发展水平来决定，而不是由它们的周边区域来决定。

这对于通过城市规模这个变量间接度量的社会组织特性，也非常适用，尽管它并没有完全解决能量获取、战争能力和信息技术的单位定义问题。对于这些特性，仍然有可能有倾向性地确定一个核心，有意地将高分和低分的地区结合在一起，人为地制造出较低的总体分数。

在《西方将主宰多久》一书中，我将下面这个问题称为“彭慕兰问题”：历史学家彭慕兰注意到，认为欧洲在工业革命以前就已经比中国

发达的历史学者们，经常试图以一种不相称的比较来证明自己的论点，即以欧洲一小块发达的核心地带——通常包括英国和低地国家——与整个中国相比较。彭慕兰指出，更相称的比较，应当是将英国与长江三角洲比较，或者将整个中国与整个欧洲相比较。

我的指数应对彭慕兰问题的主要办法是要求分析必须清楚。表2.1清楚地显示了在每个时间点上，哪个地方被认为是每个地区的核心地带。该表允许批评者们挑战定义，提出另外的核心地带，并显示出他们的修改将会如何改变分数。将假设暴露于挑战和歪曲之下，给指数建设提供了一种较好的基础，比试图制定能应对一切可能性的规矩要好。

最后，说说定量分析的问题。长达50年的关于社会进化的数值方法的辩论，其主要教训是辩论本身就跑题了。有些学者从原则上就反对定量分析的办法，另一些学者反对定性分析的办法也同样是建立在原则的基础上，但两派学者都错了。有些问题只能用定量的办法来回答，而有些问题只适用于定性的办法。如果“为什么目前西方主宰世界”这个问题果真是一个关于社会发展的问题的话，那么最好的办法是定量分析，使用一个能向我们展示需要解释的历史轮廓的社会发展指数。

The Measure of Civilization

How Social Development Decides the Fate of Nations

第三章

能量获取：交替上升

70年前，莱斯利·怀特曾说，能量获取应当是了解社会发展的基础。物质的复合性层次要想随时间的推移而维系，除非它们能自由地从所处的环境中获取能量，人类及其社会也不例外。

如果没有氧气，构成我们身体的物质的复合性层次在几分钟后就要开始瓦解：如果没有水，我们几天后就会完蛋；如果没有食物，我们至多撑上几星期，也要呜呼哀哉。如果将多个人聚在一起创造超个体，人们就必须获取更多的能量，这使得能量获取成为社会发展的基础。

我在说“能量获取”时，指的是人类全方位的能量获取，最重要的有：

食物。无论是直接消耗，还是喂养牲畜以让它们提供劳力，还是喂养肉畜以供随后食用。

燃料。无论是用于做饭、取暖、制冷、烧窑或烧炉，还是用于为机器提供能源。燃料既包括木头、煤炭、石油、天然气，也包括风能、水能和核能。

原材料。无论是用于建筑、金属制品、制陶、制衣，还是用于任何其他目的。

如此定义的能量获取与更普遍使用的生理幸福感尺度有关，但要宽泛一些。生理幸福感尺度包括诸如实际工资、人均GDP、人均GNP（国民生产总值），或者人均NDI（国民可支配收入）等。实际工资衡量校正了通货膨胀的个人收入（无论是以现金还是其他形式获取）；GDP衡量消费、生产添加的价值和一国疆域内产生的收入；GNP衡量的是加上或减去从世界其他地方转移财产收入或劳动收入而来的净收入的GDP；而NDI衡量的是GNP加上或减去以货币或其他形式从世界其他地方获取的净转移支付的GNP，包括税收和馈赠。

GDP、GNP和NDI只需各自简单地除以研究地域的人口数量，就可转换为人均数字。

经济学家通常关注的是人均实际工资、人均GDP、人均GNP和人均NDI，而不是能量获取，很大程度上是因为这些度量标准在现代经济体（即1800年后的西方、1900年后的东方和1950年后的世界其他地方）中，比种类更广泛的能量获取记录要完备得多。然而，如果在时间跨度极长、生存实践的本质变化巨大的情况下做比较，能量获取是一种更灵活的度量标准。

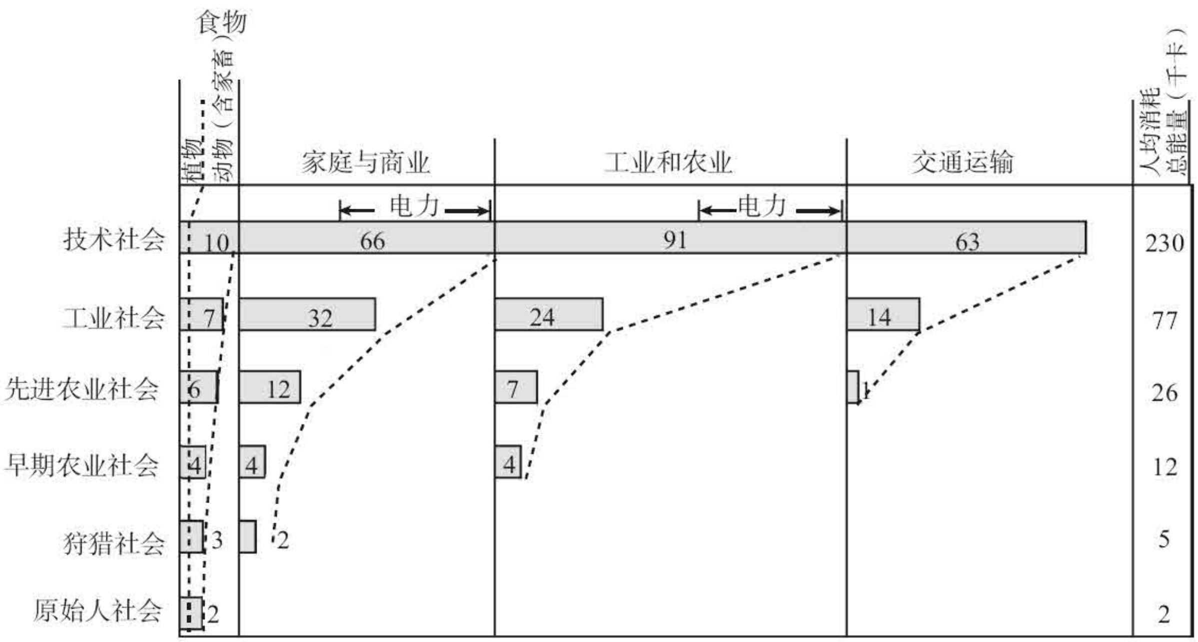
东方、西方从哪儿出发

关于人类的能量使用情况，涌现了浩如烟海的文献，贡献者有医学研究者、工程师、自然科学家、社会科学家和人道主义者。然而，相形之下，很少有人尝试对其进行历史的综合分析，而且由于不同的研究者专注于不同的能量获取特点（比如，食物消费、净能量使用、物质生活标准、总消费等），以不同的方式对其进行度量（比如，每人每天千卡消耗量、出生时预期寿命、实际工资、身高等），或者对变化进行定性分析而不是定量分析，使得即使只是形成总体概貌也变成了一个复杂的任务。因此，我首先要更严密地定义我的一些术语。

我的总体框架要从一幅广为引用的图说起（见图3.1）。这幅图最初是于1971年发表在《科学美国人》（*Scientific American*）杂志上的。在该图中，得克萨斯农工大学的地球科学家厄尔·库克（Earl Cook）对狩猎–采集社会、早期农业社会（他指的是大约公元前5000年亚洲西南部的农民）、先进农业社会（大约公元前1400年欧洲西北部的农民）、工业社会（1860年左右的西欧人），以及他本人身处的北美和西欧的技术社会的典型的人均能量获取，进行了粗略估计。库克给4项指标打了分：食物（包括喂养的家畜）、家庭和商业、工业和

农业，以及交通运输。此图表成为研究能量获取情况的历史学家一个常用的出发点。

图3.1 厄尔·库克的社会发展不同阶段能量消费（以每人每天计算）



资料来源：库克，《能量流》（*Flow of Energy*），137页。

库克将食物能量与非食物能量区别对待，这是至关重要的。人类对食物能量的消费是具有极大的强制性的：如果在一段时间内，这一数值远低于平均每人每天2 000千卡，人们就会变得虚弱而无法劳动。他们将丧失身体的机能，过早死去。然而，如果食品能量的输入在一段时间内持续高于4 000千卡，人们也会变得肥胖，患上极其严重的相关疾病，同样会有不少人过早死去。（营养学家通常会用“卡路里”来描述物理学家所说的营养“千卡”，食品包装上的营养成分表上列出的含热量实际上指的是千卡。）

食物能量的消费会随时间推移而变化，部分是因为人们会在诸如谷物等“廉价”热量食物和诸如肉之类的“昂贵”热量食物之间来回摆动（大致估算，消耗10千卡的植物才能生产1千卡的肉类）。在肉类丰富的21世纪，饮食通常会达到每人每天大约10 000千卡。然而，在非食

物的表中，能量消费变化得更显著。大多数狩猎－采集社会消耗的非食物热量都相当少：他们需要生物量用于烹饪、燃料、衣服、武器、篮子和个人饰品，但通常居住在非常简陋的栖身所里，只有极少的实用物质产品。农业社会通常有数量多得多的坚固房屋，有大量的各式各样的人工产品，而现代工业社会无疑更是能生产出数量巨大的非食物产品。在最简单的热带狩猎－采集社会，总能量获取量（食物+非食物）大致会低至每人每天4 000~5 000千卡；而在当代美国，这个数字则会高达每人每天230 000千卡，全球平均数现在大约为每人每天50 000千卡。

在历史上的大部分时期，人均非食物能量都是趋于上涨的，但人们没有多少办法将非食物热量转化为食物。结果，增加食物热量的困难成了扩大人口规模和提高生活水平的主要障碍。托马斯·马尔萨斯（Thomas Malthus）在其《人口原理》（*Essay on the Principle of Population*）中就已认识到这个问题。他写道："应当始终牢记，食物与加工产品之间存在根本区别。加工产品的原材料取之不尽、用之不竭。对这些产品的需求可以促使它们需要多少就能生产出多少。然而对食物的需求，绝对没有这样的创造力。"

甚至在史前时代，非食物能量都能稍稍缓解食物供应的压力。例如，提供肥料、改善交通以使食物从充裕的地方流动到匮乏的地方，还可以提供燃料以加工食物。然而，直到19世纪起——具有讽刺意味的是，始于马尔萨斯在世时——交通、加工、肥料和科学干预才对食物供应进行了彻底改革，使得人类的身材、寿命和健康状况得到不断改进。

尽管马尔萨斯和库克的研究都很出色，对长期经济史感兴趣的社会科学家们仍通常会忽略食物热量和非食物热量的差别，只专注于食物，他们的结论是：自一万年前农业发明到200年前工业革命之前，几乎没有发生什么变化。在最被广为引用的最近的一次讨论中，经济史

学家格里高利·克拉克（Gregory Clark）清楚地表明："（公元）1800年时世界上的人均富裕状况并不比公元前100 000年时好。"但他的论断是错误的。正如马尔萨斯所认识到的，如果良好的天气状况和先进的技术或组织促使了食物产量提高，人口往往会增长，消耗掉盈余的食物，迫使人们消费更少、更廉价的食物热量，然而尽管人均食物供应有下降的压力，从长远来看，非食物能量获取的增长，在整个全新世却得到了稳步的累积。

库克提出，虽然典型的狩猎－采集社会只能获取每人每天2 000千卡的非食物能量，但到早期农业社会时便提高到8 000千卡，而工业革命前先进的农民所获取的达到了每人每天20 000千卡。我重新构建的框架认为，从长远看（暂且忽略几个暴跌期），自大约公元前12700年冰期结束后的13 000年间，非食物能量的获取提高得较缓慢，但却很平稳，直到罗马时代的意大利——最先进的古代农业帝国的核心地带——这一数字可能会达到每人每天25 000千卡。这似乎是前工业社会可能达到的上限，相当于经济史学家E·A·里格利（E. A. Wrigley）所说的高级有机经济体和化石燃料经济体之间的界限。

在将近2 000年间，农业帝国不断冲击着这个上限，但却无法突破。到了十七八世纪，当全球化达到了能使植物和动物在大洲之间流动的程度时，投入到交通运输业的热量开始间接地转化为食物热量。然而，直到19世纪，企业家们学会了将煤燃烧释放的热能转化为动能后，非食物能量获取才得到了极大增长，使之能够转化成食物热量。这便将人类从马尔萨斯所说的陷阱中解救了出来——至少目前如此。

库克的估计当然只是个起点，因为他只提出了6个数据点（原始人社会、狩猎社会、早期农业社会、先进农业社会、工业社会、技术社会），没有尝试区分世界上的不同地区；他也没有提供他估计时使用的资料来源。因此，在重新探索西方和东方的能量获取时，我先将库克的数据作为出发点，为给定体制内"正常"的消费建立一个数量级，

然后利用更详细的证据，估算出每个时间点上东方和西方的核心地带实际上距离这些正常的数据有多远。

度量的单位

在书中，我使用下列传统的度量单位及缩写：

1卡=将1立方厘米的水的温度提升1℃所需要的热量

1卡= 4.2焦耳

1焦耳= 0.238卡

1英制热量单位= 1 055焦耳

1吨小麦当量= 3 300 000千卡

1吨原油当量 = 10 038 000千卡

1公升小麦当量= 0.78千克小麦当量=2 574千卡

1兆焦=239 999千卡

1瓦特= 1焦耳/秒

1马力= 750瓦特

成人基本生理需求所需食物热量=2 000~2 700千卡（=8~11兆焦≈90瓦特）/（人·天）

证据的本质

关于能量获取的可靠的统计资料，在东方的核心地带，只能部分回溯到20世纪，在西方也只能回溯到19世纪早期。即便如此，这些资料通常也都会遗漏大量的农业社会用于燃料和建筑的生物量。零星的

统计资料在中国和日本的部分地区可追溯至19世纪，在西欧可至少追溯至17世纪。在此之前，无论东方还是西方，都只有文本记录，偶尔才有一些计量文献，在中国可回溯至公元前1200年，在美索不达米亚和埃及可回溯至公元前3000年，但这些资料都无法提供像现代时期那样可供了解的细节。

我们在时间上回溯得越远，就越必须依赖于考古证据和比较证据。前者有时能为我们勾勒出一幅关于谷物和技术的非常清晰的画面，以及关于贸易水平和生活水准的一种尽管模糊但仍然很重要的感觉。结合现代语境中记录良好的关于类似的谷物、技术、贸易和生活方式的能量产额的比较证据，我们可以大致了解能量获取。我们偶尔也可以对照一些完全不同类别的证据，如冰蕊和泥炭沼的污染记录等，反复核查结果。

将这样形形色色的资料结合在一起，当然是个挑战，需要时常施展猜测的功夫。一方面，这使得专家们对于公元1900年之前的东方和公元1700年之前的西方的精确分数，永远不可能取得一致意见；但另一方面，证据的确形成了没有专家会质疑的历史上能量获取的参数。例如，没人会认为公元1000年西方的核心地带（大约在伊拉克－埃及一带）或东方的核心地带（黄河流域）的能量获取，会像1 000年后的美国或日本一样高，或者就此而言，像1900年、1800年，甚至1700年时的核心地带那样高。同样的，也不大可能会有专家提出公元1000年时西方的能量获取会和1 000年前罗马帝国时一样高，但几乎所有专家都会同意，那时的能量获取会比大约公元前1000年“黑暗时代”的地中海地区高。在东方，大多数中国经济史学家也许会同意，东方的能量获取水平在公元1000年左右的宋朝，比在公元1年左右的汉朝要高，比在公元前1000年左右的西周时期更是要高得多。任何违背这些看法的结论，都会招致严密的审查。

在一定限度内，我们当然可以建立一个大致的、约略的能量消费数字，但更重要的是，我们能否将误差幅度控制到足够小，从而大概估计能量获取状况，使我们能够确认对于为什么目前西方主宰世界这个问题的最好解释，究竟是长期注定论，还是短期偶然论，抑或根本是某种其他理论。

中断与增长：西方能量获取

计算不同时期的能量获取的最好方法，是从最广为人知的时期向最不为人知的时期推算，因此我不是从公元前14000年起始，不断向前推进至公元2000年，而是让我的讨论从目前开始，回溯至公元1700年，然后向后进行两次大跳跃，再去填补三个时期之间的沟壑。第一个跳跃是跳回到约公元前500~前200年的古典地中海世界，有好几位经济史学家新近提出了一些消费水平的数据；第二个跳跃是跳回到我们故事的开端，大约公元前14000年时，在这个时间点上（尽管会令非考古学家们大吃一惊），我们可以对冰期晚期狩猎－采集社会的消费，做出非常自信的估计。

表3.1 公元前14000~公元2000年西方能量获取情况

年代	千卡/（人·天）	分数（分）	年代	千卡/（人·天）	分数（分）
公元前14000年	4 000	4.36	公元前500年	23 000	25.06
公元前13000年	4 000	4.36	公元前400年	24 000	26.15
公元前12000年	4 500	4.90	公元前300年	26 000	28.33
公元前11000年	5 000	5.45	公元前200年	27 000	29.42
公元前10000年	5 000	5.45	公元前100年	29 000	31.06
公元前9000年	5 500	5.99	公元前1年	31 000	33.78
公元前8000年	6 000	6.54	公元100年	31 000	33.78
公元前7000年	6 500	7.08	公元200年	30 000	32.69
公元前6000年	7 000	7.63	公元300年	29 000	31.60

公元前5000年	8 000	8.72	公元400年	28 500	31.06
公元前4000年	10 000	10.90	公元500年	28 000	30.51
公元前3500年	11 000	11.99	公元600年	26 000	28.33
公元前3000年	12 000	13.08	公元700年	25 000	27.24
公元前2500年	14 000	15.26	公元800年	25 000	27.24
公元前2250年	16 000	17.44	公元900年	25 000	27.24
公元前2000年	17 000	18.52	公元1000年	26 000	28.33
公元前1750年	19 000	20.65	公元1100年	26 000	28.33
公元前1500年	20 500	22.34	公元1200年	26 500	28.88
公元前1400年	21 000	22.88	公元1300年	27 000	29.42
公元前1300年	21 500	23.43	公元1400年	26 000	28.33
公元前1200年	21 000	22.88	公元1500年	27 000	29.42
公元前1100年	20 500	22.34	公元1600年	29 000	31.06
公元前1000年	20 000	21.79	公元1700年	32 000	34.87
公元前900年	20 500	22.34	公元1800年	38 000	41.41
公元前800年	21 000	22.88	公元1900年	92 000	100.00
公元前700年	21 500	23.43	公元2000年	230 000	250.00
公元前600年	22 000	23.97			

图3.2 公元前14000~公元2000年西方能量获取情况（线性－线性标尺）

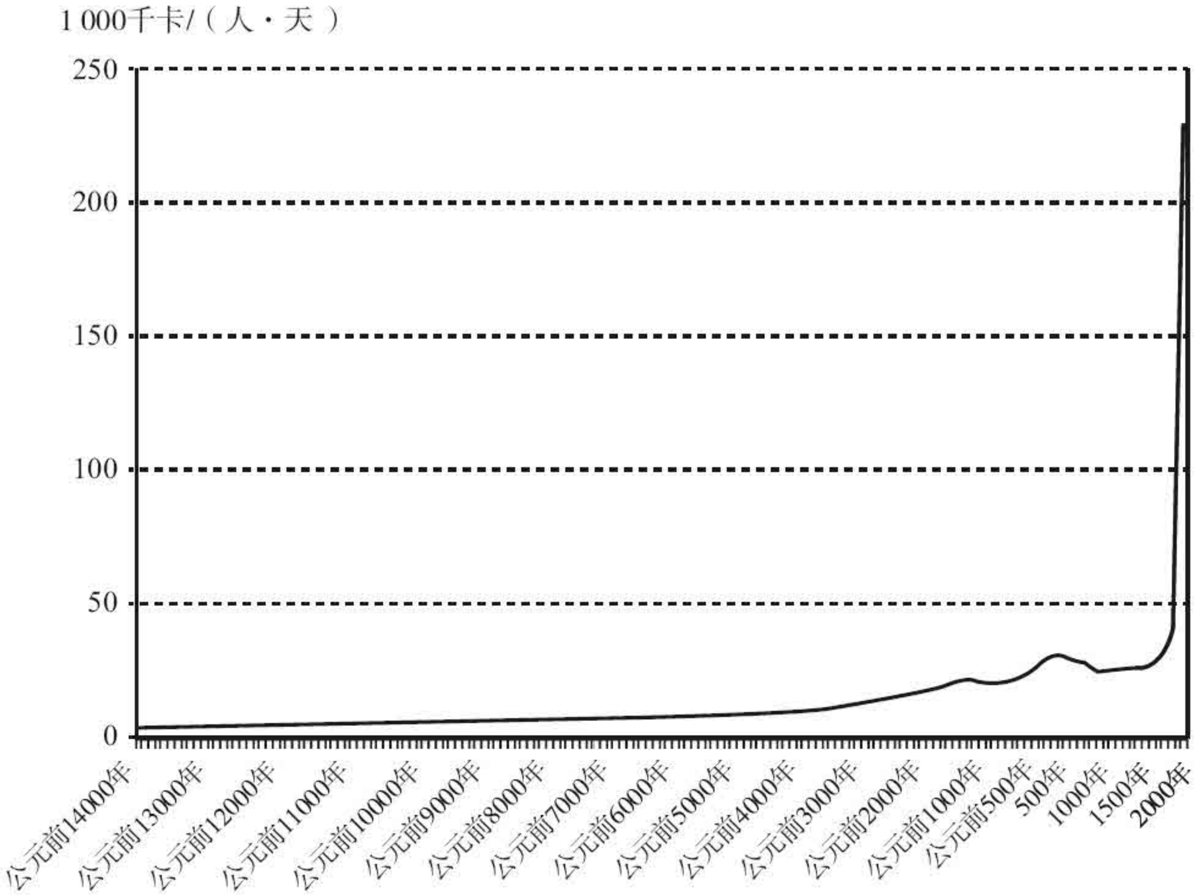
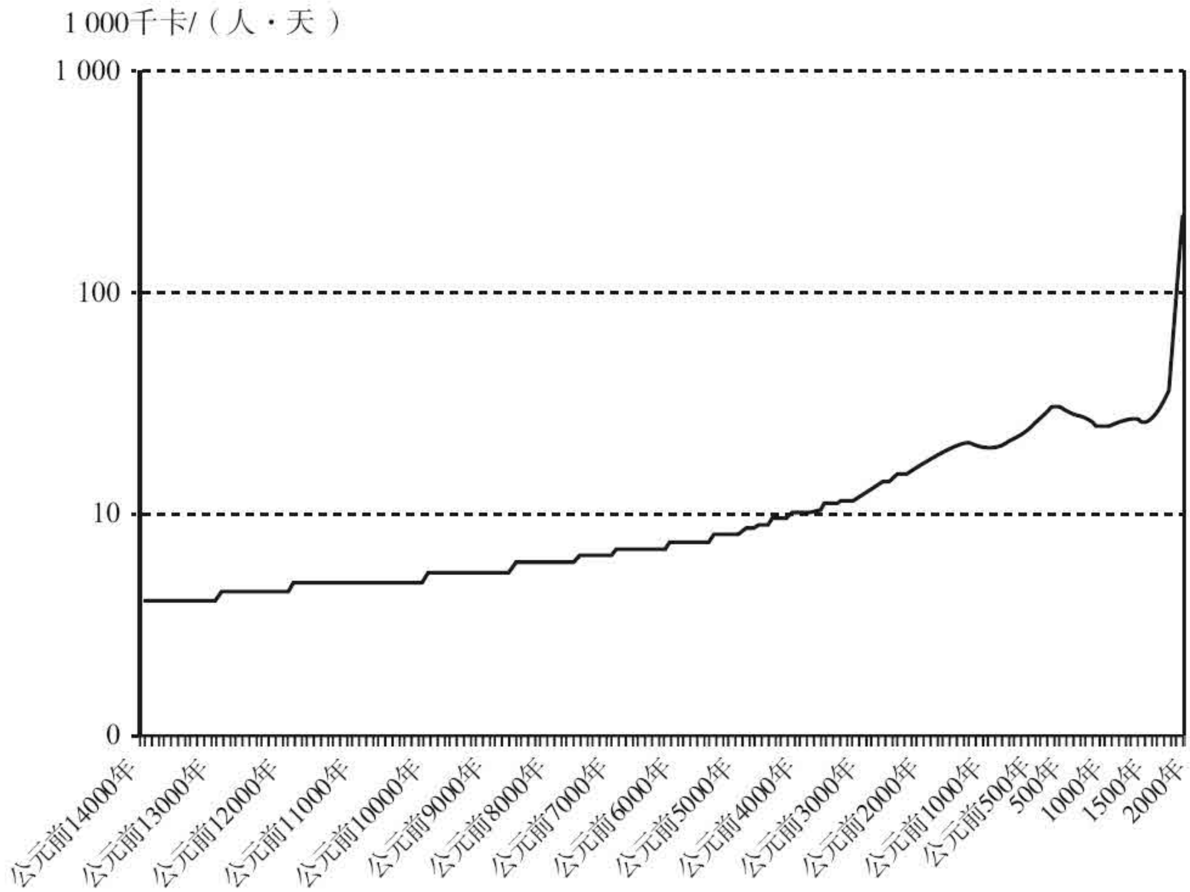


图3.3 公元前14000~公元2000年西方能量获取情况(对数–线性标尺)



不远的过去(1700~2000年)

公元2000年的统计数据质量很好，据此可以得出西方核心地带（美国）的人均获取食物+非食物总能量为（每人每天）大约230 000千卡。按照第二章中介绍的方法，每人每天230 000千卡——史上最高的能量获取水平——可得到满分250分，意味着在指数中，每人每天920千卡得1分。

我们关于1900年甚至1800年西方最先进的经济体（大西洋北岸一带）的资料，至少在某些方面是相当完备的。有可回溯至1700年的关于欧洲某些部分的工业产值的比较丰富的资料，但主要难题是怎样将这个信息与用作燃料、房屋和衣服等的生物量结合起来。最依赖于生物量的农民往往不会留下太多的文字记录，这就迫使我们关注那些根据比较证据所做的估计，再与文学和艺术作品中的定性证据进行对比

核实。定性证据通常极其丰富，但又必须将不同来源的资料汇集起来，这不可避免地会扩大误差。

将化石燃料和生物燃料的数字，与安格斯·麦迪森（Angus Maddison）的人口数据相结合，显示西方核心地带典型的能量获取数字为：1900年，每人每天约92 000千卡；1800年，每人每天约38 000千卡。据我的大致估计，1900年的每人每天92 000千卡可分解为约41 000千卡来自化石燃料，8 000千卡来自食物/牲畜饲料，43 000千卡来自非食物生物量；而1800年的每人每天38 000千卡，可分解为约7 000千卡来自化石燃料，6 000千卡来自食物/牲畜饲料，25 000千卡来自非食物生物量。1900年的每人每天92 000千卡和1800年的每人每天38 000千卡这两个数字，都与库克所估计的1860年的先进西方经济体为每人每天77 000千卡的数字完全相容，在这样的数据可等同的范围内，与遗嘱检验记录和工业考古学所提供的关于家庭用品增长的证据，看来是一致的。1800年和1900年的数字比2000年的数字所容纳的误差幅度要大，但与历史文献中关于能量使用给人们留下的深刻印象，以及罗伯特·艾伦（Robert Allen）所重新确定的实际工资的变化趋势，也是相符的。

我所估计的西方核心地带的人均能量获取在1800~1900年有242%的增长，低于已被广泛接受的发达的欧美核心地带工业产值增长的统计数字。那是因为在估算工业产值时，生物量和肌肉力量通常完全不纳入计算，使得总体能量获取的情形被扭曲了。19世纪工业产值的一个重要剖面是在向着取代生物量和肌肉力量，而不是增加它们的方向发展，在这一过程中要容许工业核心地带有比以往高得多的人口密度，而又不至于产生环境灾难。

当我们回顾1800年之前时，不确定性无疑大大增加了，但我们的估计仍然有强大的约束力。西方的能量获取在18世纪比在19世纪显然要增长得慢，但比在17世纪或16世纪则要增长得快；如果库克的推断

——中世纪晚期先进的农业社会已能获取每人每天26 000千卡的能量——是正确的，1700年左右早期现代化的西北欧所消费的能量，一定在每人每天30 000~35 000千卡。

1700年和1400年西方核心地带能量获取的比率大致为5∶4，这一猜测是基于大量的文献资料和考古证据做出的，内容涵盖整个社会生活的方方面面，如房屋质量的改善，家庭用品品种的增多和质量的改善，西北欧地区实际工资的上涨，昂贵热量消费的增长，工作时间的增加，等等。

安格斯·麦迪森估计西欧的人均GDP在1500~1700年，从798美元（以国际元来表示，这是一种与1990年的1美元具有同样购买力的假设单位）增长至1 032美元。有好几位经济学家都认为麦迪森的数字低估了，但其总体趋势似乎没有错误——只要我们记住几乎所有的计算结果似乎计算的只是非食物热量。成人的身高是反映儿童营养水平的强有力的指标，但在1700年与在1400年似乎变化不大。

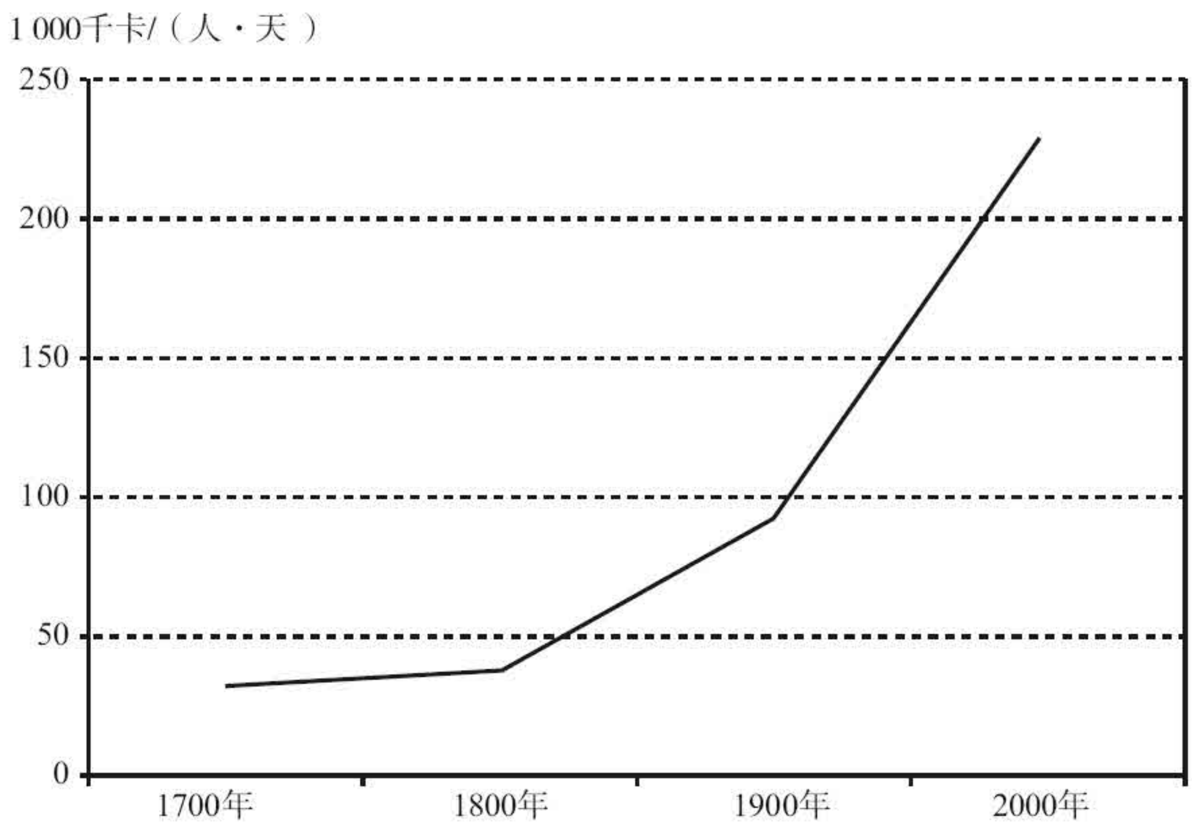
我的数字——1700年为32 000千卡——当然是个推测，但我认为误差不超过10%，原因如下：

一是，如果西北欧的消费在1700年时已经超过了每人每天35 000千卡，但是到1800年时只增长到38 000千卡，就很难解释工业和运输业上消耗的所有额外的能量从何而来（正如经济学家罗伯特·艾伦所展示的，实际工资在1750~1800年也许是下降了，随后也增长得很缓慢，直到1830年，因为新的经济精英获得了大量利润，又进行了再投资）。

二是，西北欧的能量消费在1400年时已经达到了每人每天26 000千卡，假如在1700年仍然低于每人每天30 000千卡，就很难解释贸易、工业、农业和林业在15~17世纪，怎么可能像我们所知的那样迅猛发展，而能量获取增长得却这样缓慢。

三是，假设西方的能量消费在1400年后迅猛上升，在1700年左右却仍低于每人每天30 000千卡，在此情况下，我们可以将1400年的数字从每人每天26 000千卡压低至每人每天20 000千卡，那么我们将不得不或者承认，（按前现代化的标准）生产力极高的1400年左右的欧洲社会，在能量获取方面还不及大约3 000年前青铜时代的地中海东南地区，这似乎不大可能；或者承认，大约公元前1600年的能量获取还要低，或许在每人每天15 000千卡左右，这反过来又要求我们将更早的数字压得更低。由于我们可以为后冰期时代的能量获取设一个至少每人每天4 000千卡的下限，将公元前第二个千年的能量获取水平压低到每人每天15 000千卡，这么一来，像公元前1500年左右乌尔这样的考古遗址有坚固的房屋，而像公元前12000年左右以色列的恩马拉哈这样的遗址只有非常简陋的栖身所，两者之间生活水准存在的巨大差异就难以解释了。

图3.4 1700~2000年西方能量获取情况



古典时期（公元前500~前200年）

过去几年，有几位历史学家和经济学家曾尝试量化古典时期地中海地区的实际工资和人均GDP。这些计算虽然与本书所定义的能量获取不是一回事，但是迈出了非常有益的一步。

实际工资

我们有关于古代地中海地区工资和食物价格的信息，参差但有用，据此也能计算出部分时期和地方某类人每天能买得起的小麦的数量。在最近的一篇重要文章中，沃尔特·沙伊德尔（Walter Scheidel）循早期现代史学家扬·范·岑登（Jan Van Zanden）之例，将古代的工资数据换算为“小麦工资”，借以表示一名工人一天收入可以买到的小麦的公升数。掌握了这样的信息，以及1公升小麦（0.78千克）含2 574千卡，我们就能计算出由工资水平代表的能量获取情况了。

沙伊德尔证明，在公元前400年之前不久，一名雅典成年男子的实际工资，每天可以购买含有超过22 400千卡能量的小麦，而到了公元前3世纪20年代，实际工资的购买力上升到相当于每天33 500~40 000千卡。这是极高的数字了，接近于18世纪甚至19世纪早期的西方核心地带了。

沙伊德尔关于公元前几个世纪罗马时代的意大利的数据，变化要大得多。其中罗马城的工资居于相当于每天15 500~43 000多千卡，而庞贝的每日工资是12 000~30 000千卡。这些数字的平均值为每天大约25 000千卡，但正如沙伊德尔所指出的，考虑到变异量，很难过多相信这个数字。

这些数字代表着前进的一大步，但在将实际工资与能量获取相关联的道路上，也有后退的两步。首先，正如沙伊德尔本人所强调的，数据点如此分散，我们无法判断数据是否典型。这种情况只有一例，

在古代欧亚大陆的西部，即公元前385~前61年的巴比伦，我们知道许多商品的真正详细的一系列价格，然而价格波动得也很剧烈。由于我们经常不得不面对一连几个世纪无资料的情况，只搜集到单独的价格点，我们很可能会被零星信息误导。

其次，工资水平与总体的食物+非食物能量获取之间的关系很不明朗。我们只有少量行业的工资信息，许多人也许部分或很大程度上是在非货币化经济中从业的，他们在家庭农场或家庭商铺中谋生。在古典时期的雅典，工资数据是由诸如军饷和担任公职的俸禄之类的国家雇佣主宰的。在这些部门，国家担任了独家垄断买方的角色，这就使得工资水平与私有领域的关系变得很难琢磨了。

罗马的数据没有如此严重地偏斜于国家付酬，但也有它们自己的问题。我们不知道无记录的行业和有记录的行业的对比情况如何，也不知道家庭通常需要通过什么样的收入来源来补充文献中提到的工资收入，或者典型的家庭在能量获取方面有多少来自完全在货币化经济之外的生物量。

人均GDP

还有一个办法是计算古代社会的GDP，再除以其人口规模。有多位历史学家和经济学家提供了对公元最初两个世纪罗马帝国的估计（见表3.2）。这个办法避免了实际工资的一些问题，但也产生了其自身的一些新难题，最明显的是计算须依赖于一系列消费情况。沙伊德尔和弗里森（Friesen）甚至承认“研究罗马世界的学者们，如果不熟悉我们的方法，也许会倾向于把这种项目贬斥为一团乱麻般的臆想”。

表3.2 罗马人均GDP估计

	千克小麦当量（人·年）	千卡（人·年）	千卡（人·天）
霍普金斯	491	1 620 000	4 438

麦迪森、戈德史密斯	843	2 780 000	7 616 仅算意大利：12 712
特明	614	2 030 000	5 561
麦迪森、戈德史密斯 （经沙伊德尔、弗里森的数据调整）	620	2 050 000	5 616 仅算意大利：9 370
埃及“概况” （沙伊德尔、弗里森）	390	1 290 000	3 534
埃及“可观数字” （沙伊德尔、弗里森）	940	3 100 000	8 493
沙伊德尔、弗里森	714	2 360 000	10 710
301年，戴克里先价格法令（根据艾伦的资料）	204	670 000	1 836

资料来源：霍普金斯，《罗马帝国的税收和贸易》(*Taxes and Trade in the Roman Empire*）；戈德史密斯，《规模和结构的估计》(*Estimate of the Size and Structure*）；麦迪森，《世界经济的轮廓》(*Contours of the World Economy*）；特明，《估算GDP》(*Estimating GDP*）；沙伊德尔和弗里森，《经济的规模》(*Size of the Economy*）；艾伦，《罗马究竟有多繁荣？》(*How Prosperous Were the Romans*？）

最重要的假设是估计最小食物需求量，这一“提高”能表现出非食物消费情况；另一个假设是要表现出政府开支情况，还要推测出每年工作日的典型数字。关于所有这些数据，意见都很不一致。公元最初两个世纪人均GDP的估计结果，低至古代史学家基思·霍普金斯（Keith Hopkins）提出的相当于每人每工作日7 364千卡，高达经济学家雷蒙德·戈德史密斯（Raymond Goldsmith）和安格斯·麦迪森提出的每人每工作日12 636千卡。沙伊德尔和弗里森强调研究一系列估计的必要性，但他们的确提出了每人每工作日10 710千卡，作为总结性的数字（17 000万人口在220个工作日中生产了5 000万吨小麦当量）。将估算方法与罗马帝国治下的埃及的资料相结合后，他们认为实际的数字必然在每人每工作日5 864~14 091千卡，好几种不同的方法也都会合于这一范围。

这些能量获取的分数比从实际工资推算出来的要低得多，这似乎出于多个原因。原因之一是，人均GDP的估算法是运用于整个罗马帝国，而不是意大利核心地带的。这又一次提出了关于单位选择的“彭慕兰问题”。我们需要专注于西方最发达的核心地带，在此应该是意大利。麦迪森意识到了这一点，提出流入意大利的税收和贡赋提高了其NDI，使之比帝国其他地区高出2/3，这将使麦迪森对意大利能量消费的估算被推高到每人每工作日12 712千卡（或者，使用沙伊德尔和弗里森对他的分数提出的调整，每人每工作日9 370千卡）。

然而，意大利的这个分数，甚至仍然低于沙伊德尔所使用的从罗马和庞贝的实际工资推算出的能量获取数字范围的最低值，接近于库克计算的早期农业社会（他指的是大约公元前5000年亚洲西南部的农业社会）的分数。对此的解释是，用于所有对GDP数字的“提高”，非常严重地低估了罗马经济中用于燃料和建筑、风能和水能，以及原材料的生物量。霍普金斯只提高了33%，试图覆盖种子和损耗，即使戈德史密斯给出的最高估计（得到了麦迪森、沙伊德尔和弗里森的同意）也只有75%。关于能量获取的对比数据表明，真实水平一定高得多。

瓦茨拉夫·斯米尔（Vaclav Smil）在其对生物质能量精妙的研究中，以能量密度为标准将生物质燃料分成了两类（见表3.3）。他的极低密度类（泥炭、生材、草类）每千克产生5~10兆焦（相当于1 200~2 400千卡），而低密度类（作物残茬、风干木材）每千克产生12~15兆焦（相当于2 880~3 600千卡），看来与古罗马最为相关。煤炭的使用在罗马帝国并非无足轻重，尤其是在北方省份，但化石燃料的确还不是主要的燃料来源。

表3.3 能量密度

密度	能量	能量密度（每千克兆焦）
食品		

极低	蔬菜、水果	0.8~2.5
低	薯类、牛奶	2.5~5.0
中等	肉类	5.0~12.0
高	谷物和豆类	12.0~15.0
极高	食油、动物脂肪	25.0~35.0
燃料		
极低	泥炭、生材、草类	5.0~10.0
低	作物残茬、风干木材	12.0~15.0
中等	烟煤	18.0~25.0
高	木炭、无烟煤	28.0~32.0
极高	原油	40.0~44.0

资料来源：斯米尔，《大众能量学》（*General Energetics*）

我们当然没有罗马帝国生物质燃料使用情况的统计数字，但我们的确有一些具有提示性的比较证据。20世纪热带狩猎－采集群体通常靠每人每年不到500千克的生物质燃料就能过活，其中大部分估计是极低密度类型的，也就是说，其能量获取量大概为每人每天1 300~2 600千卡。气候较冷地区的农业社会化通常使用的生物质燃料约为每人每年2.5吨，估计既有低密度类的，也有极低密度类的；假设低密度类/极低密度类的比率为50∶50，那么能量获取量约为每人每天12 329~22 191千卡。18世纪西北欧和北美先进的有机经济体的使用量为每人每年3~6吨。如果我们再次假设低密度类和极低密度类燃料的比率为50∶50，那么能量获取量将为每人每天21 699~43 397千卡。

这些关于其他社会生物质燃料使用情况的数据，与库克对西欧中世纪晚期先进农业社会的非食品能量消费的估计——每人每天20 000千卡——是一致的。最重要的问题是，古代地中海地区的经济在这个范围内需要摆在什么位置。要回答这个问题，我们必须求助于考古学。

考古证据

考古方法包括调查古人为获取能量所留下的实际物质残存，其形式包括人类和动物的骸骨、碳化的种子、花粉、房屋、人工产品和污染留下的化学痕迹。这种实地考察的办法比更程式化的实际工资法和人均GDP法麻烦得多，但更具有实证效力。最重要的是，比之非常抽象的人均GDP法，这能描绘更清晰的画面，并显示出无论实际工资法还是人均GDP法，都严重地低估了古代社会的能量获取能力。

考古证据证实了实际工资法产生的数据给人们留下的印象。按照古代社会的标准，公元前4世纪的希腊人获取能量的能力很高。他们的饮食相对较好，肉类的含量通常较低，尽管地方与地方之间差别较大。橄榄油、葡萄酒、水果、大蒜和鱼类占了相当大的比重，尽管鱼类的消费量像肉类一样，一地与另一地之间差别很大。食物消费不足以将成年男子的平均身高提高到168厘米以上，但鉴于“昂贵的”热量的数量，按照地中海地区古代社会的标准，典型的希腊人摄入的食物能量一定是相当高的，也许能达到每人每天4 000~5 000千卡。

典型希腊人的饮食好（人口增长也快），部分也许是因为太阳活动减弱，使得公元前800年气候从亚北方气候转为了亚大西洋气候，给地中海地区带来了更凉爽、更湿润的天气，有利于依赖冬季的雨水种干谷的农民。然而，最近对地中海东部地区的80份论文的综合研究，显示出了超乎寻常的区域差异，而在大约公元前800~前200年，只有细微的变化。

无论气候扮演着什么角色，希腊人的行为变化似乎的确发挥着作用。自20世纪80年代起，实地调查的考古学家们就认识到，对希腊农业模式的旧看法，即认为其效率低下和规避风险，绝对是不正确的，因为这样的农业体系不可能生产出足够的粮食，来供养古希腊世界的人口密度。

关于聚落形态和发掘出来的农庄的证据，表明公元前500~前200年可能出现了邻近土地上形成集中劳动群体的变化，使得肥料得到大量施用，通常生产是为了适应市场，通过干谷耕作获得的收益至少要到19世纪才能与之媲美。孢粉数据支持了这一观点，谷物和橄榄生产的巅峰时期，在大约公元前500~前200年，不仅出现在希腊，而且遍及地中海东部地区，甚至深入到亚洲的伊朗西部。

古典时期希腊的房屋又大又舒适，通常都有240~320平方米的屋顶空间。关于房屋价格的资料存在争议，但一般来说一座房子的成本大概是1 500~3 000德拉克马（古希腊银币），当时5 000千卡的日常饮食需要约半个德拉克马——这意味着一座普通房屋的成本相当于1 500万~3 000万千卡。按30年的房屋寿命期平摊，相当于每天将近1 375~2 750千卡。（我们无从得知希腊人预期的房屋寿命是多久，但30年似乎大致符合从考古遗址观测到的房屋重建率。）

更难量化的是窑炉、熔炉、作坊等代表的人均能量消费。正是这些窑炉、作坊生产了我们在希腊人的房屋中，在神庙、堡垒、武器和盔甲、战船、公共建筑、私人纪念碑、道路、港口、艺术作品，以及考古学家们发掘的无数其他目标旁发现的所有人工制品。还有从远到乌克兰和埃及等地的农场运来的大量食物，其运输成本所代表的人均能量消费。然而，对比古希腊定居点（如毁于公元前348年，如今已被发掘得很详细的奥林索斯）与中世纪和早期现代北欧地区的定居点（如英国的沃勒姆－珀西）的房屋质量和手工艺品的丰富程度，更不用说对比中世纪和早期现代的希腊，会给人们留下深刻的印象：古希腊人享受着极高的物质生活水平。

同样惊人的还有，古希腊人不仅承受得起相当高水平的非食物消费，并且在公元前4世纪的爱琴海一带也承受得起较高的人口密度。在希腊的好几个部分，公元前4世纪的人口密度都无法望其项背，直到20世纪。如此多的希腊人居住在城镇或小城市里，而不是住在小村或农

庄中，这个事实一定意味着他们的能量获取达到了非同寻常的高度。杰夫·克朗（Geof Kron）在一篇重要论文中利用住房证据，论证了在许多方面，普通的希腊人实际上过着比18世纪普通的英国人都要好的生活。

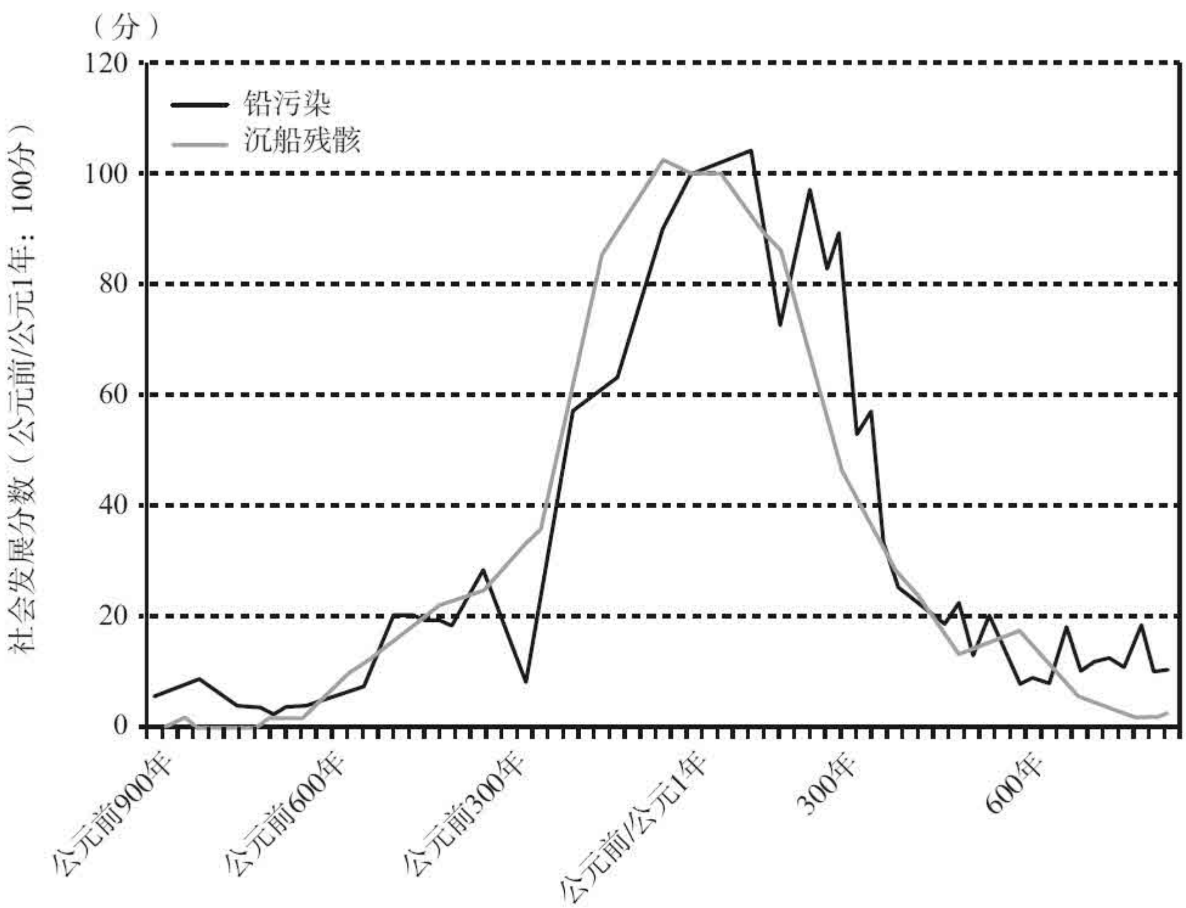
希腊考古资料清晰地指向高能量获取量（按照前现代的标准）。我估计公元前4世纪的这个数字在每人每天20 000~25 000千卡（更可能接近于较高的数字而不是较低的数字），是从“黑暗时代”的水平——在公元前1000~前800年，每人每天接近于16 000千卡——突然跃升上来的。

大量关于罗马证据表明，公元1~2世纪意大利地区的能量获取，比公元前4世纪时的希腊还要高。农业产量的水平仍然存在争议，尽管按照前现代的标准，埃及的灌溉农业的产量似乎极高。对于消费的定量研究——包括从定居点发现的牲畜骸骨，到船只残骸的数量、因工业活动而产生的铅和锡污染的水平、森林采伐的程度、公开铭刻石头的频率、流通中的钱币数量，以及沿德国边境发现的考古出土物的数量在内的一切——也指向了同样的结果：地中海的人均获取能量在公元前第一个千年增长迅猛，其巅峰时期是公元前100~公元200年，然后在公元第一个千年的中期出现下跌。图3.5展示了船只残骸数量的升降（通常被视为海上贸易水平的表现），与西班牙潘尼柳–韦柳地区年代确定的沉积物中铅污染水平的高低密切匹配。

每类资料都有其自身的局限，但没有一个论点能令人信服地解释清楚公元前第一个千年非食物消费有很惊人的增长，以及为何其巅峰出现在公元头两个世纪。船只残骸的资料和罗马城周边的运输用陶器的巨大垃圾堆（其中仅泰斯塔西奥山就包含2 500万个陶器的残片，这些陶器曾被用于装船运输两亿加仑橄榄油），也证明了非食物能量被用于增加食物供给，而且“昂贵”食物热量的消费水平异乎寻常地高。一些学者还探明了公元头两个世纪人们身高有所增长，尽管另外一些

学者比较悲观，认为罗马帝国早期意大利成年男子的身高普遍不到165厘米，这使得他们比铁器时代或中世纪的意大利人要矮。更多的证据，以及更多的与之相应的统计技术的运用，应当能解决这个问题，我们必须寄希望于格尔杰·克莱恩–戈尔德维克（Geertje Klein-Goldewijk）的罗马人骨骼数据库的面世。

图3.5 船只残骸和铅污染情况所显示的公元前后第一个千年经济增长和下跌情况



资料来源：帕克，《古代船只残骸》（*Ancient Shipwrecks*）；许兰德等，《精炼前工业化时代》(*Refining the Preindustrial*）

像在希腊的情况一样，房屋证据的信息量也许是最大的，罗伯特·斯蒂芬（Robert Stephan）和杰夫·克朗目前正在搜集和分析这方面的资料。来自埃及和意大利的资料已经表明，公元初的几个世纪，典型的

罗马房屋甚至比古希腊的房屋还要大，而且（按照前现代的标准）复杂精细的水管设施、排水系统、屋顶和地基遍及社会各阶层的房屋。

罗马考古遗址物质产品的激增甚至更令人惊讶。轮制的、烧制得很精美的陶器，装葡萄酒和橄榄油的双耳瓶，还有贱金属的饰物和工具的大规模生产，在公元头几个世纪达到了前所未有的水准。分布图也显示，到公元200年时，贸易网络之广大和密集，远超后世，至少要到17世纪才又恢复到那时的水准。与帝国正式边界之外很远的印度的贸易规模，尤其令人印象深刻。

考古数据表明，通过考察实际工资，尤其是人均GDP的办法来研究罗马经济，会低估罗马核心地带的能量使用情况。迄今为止，所有的人均GDP计算均以人类对食物热量的生理需求为出发点，对非食物消费进行了随意的“提高”，既没有考虑生物质能量的比较证据，也没有考虑非食物消费异乎寻常的激增的考古证据。正如前面所提到的，至今出现的最大的“提高”为75%，但比较证据表明，对于复杂的农业经济来说，这个数字甚至都太低了。

库克总结说，即使对于一个“标准的”先进农业经济体，“提高”的幅度都应在300%以上。考古证据很清楚，大约公元前200~公元200年的罗马时代的意大利，绝非一个“标准的”先进农业经济体。目前还没有办法非常精确地确定应当对其“提高”多少，但考古证据给我的启示是，应当比古希腊高出许多，很可能应提高400%以上。因此，公元1世纪时罗马核心地带总能量获取能力大约为每人每天31 000千卡。

这一估计使得公元100年左右的罗马核心地带的能量获取能力，仅比公元1700年的西北欧核心地带稍稍落后一点。

对罗马经济来说，这是一个比人均GDP的估计更乐观的评定，但能解决对罗马经济观察方法不一致的问题。麦迪森的数字显示公元初几个世纪的罗马帝国，与1500年的西北欧最相当，不过他随后又指

出，罗马的城市化水平实际上与约1700年的西欧更接近，而不是约1500年。虽然沙伊德尔和弗里森也曾总结说，公元2世纪时整个罗马帝国范围的经济不像1580~1600年的荷兰或1680~1700年的英国的经济那样复杂、精细，但他们的确提到过，意大利核心地带的表现要好得多。经济学家保罗·马拉尼马（Paolo Malanima）也得出了同样的结论。

我只知道另外还有两个人，曾试图用我在这里使用的术语，计算整个罗马帝国的能量获取。第一次讨论是瓦茨拉夫·斯米尔在其《为什么美国不是新罗马》（*Why America Is Not a New Rome*）一书中提出的。这本书旨在彰显当代美国和古罗马之间的不同。斯米尔非常正确地强调的一个观点是，两者在能量获取方面存在巨大的鸿沟。然而，斯米尔在试图证明这个非常有效的观点时，所提供的罗马人能量使用数据是难以置信地过分低估了。他认为当代美国人使用的能量比罗马人高出30~50倍，这将使罗马人的总能量获取数值定为每人每天4 600~7 700千卡。如果我们假设其中大约2 000千卡是食物（这意味着忽略考古证据所显示的来自肉类、食用油和葡萄酒等的相当高的昂贵热量消费水平），那么就只剩下每人每天2 600~5 700千卡来涵盖所有其他能量消费了。为了证明这个估计是正确的，斯米尔提出罗马人使用的燃料每人每年只有相当于180~200公斤的木材，或者说大致为每人每天1 750~2 000千卡的热量。

这些数字与关于罗马人的消费状况，或罗马时代沼泽、冰蕊和湖床的铅污染程度的考古证据不相符。斯米尔的数字与他在《世界历史的能量》（*Energy in World History*）一书中使用的前现代时期生物量的数据，也不兼容。斯米尔对罗马的估计，使其能量获取能力与一些有记录的最简单的农业社会为伍。我本人的估计与洛·卡西欧（Lo Cascio）和马拉尼马的计算大致相当，我们都将巅峰时代的罗马（约公元100年）的能量获取能力与1700年的西北欧相提并论，而麦迪森和沙伊德尔、弗里森则认为其与16世纪的西北欧相近。然而，斯米尔在

《为什么美国不是新罗马》一书中提出罗马人获取的非食物能量只有每人每天2 600~5 700千卡，这就将罗马人的水平降至不足斯米尔本人在《世界历史的能量》一书中估计的18世纪的西北欧能量获取数字（每人每天21 700~43 400千卡）的1/8了，使罗马人更接近于狩猎－采集社会，而不是早期现代农业社会。所有其他类的证据都使之显得实在太低了。

第二个讨论是保罗·马拉尼马在他的论文《罗马世界的能量消费和能量危机》（*Energy Consumption and Energy Crisis in the Roman World*）中提出的。这篇论文是其在2011年罗马美国学院的一次会议上发表的。该论文的附录之一直接回应了我在《西方将主宰多久》一书中的论点，提出罗马人获取的能量在巅峰时期为每人每天6 000~11 000千卡。这大致是斯米尔所估计的数字的两倍，却不及我的数字的1/3。

我们的计算之间的某些差异是明显的。正如上文所述，不同类型的食物能量有不同的成本；通常要耗费大约10千卡的饲料才能产生1千卡的肉食，这意味着向肉类消费转变的时期，也是人均能量消费增长的时期。以面包和水为食的人，也许会像以牛排和香槟酒为食的人一样，吃掉同样千卡数量的食物能量，但食用牛排/香槟酒代表着高得多的总体能量消费水平。考古证据表明，罗马时期大多数人的饮食成本都有巨大增长。这在罗马本地最为显著。葡萄酒和橄榄油消费的爆炸性增长创造了泰斯塔西奥山，但即使在最简陋的村落遗址，也出现了向更昂贵的食物热量转化的惊人证据，覆盖的人口达数千万。虽然普通罗马人的饮食并非牛排和香槟酒，但的确他们至少有橄榄油和进口葡萄酒。

马拉尼马将用于建筑、工业和运输的材料排除在了能量内容之外。在罗马时代之前的大部分时间，这种定义上的差异不会对计算产生重大的影响，因为那时候的建筑、工业和运输业一直非常简单。但

考古证据又一次清晰地显示，罗马时代与前罗马时代最巨大的反差之一就是所有这些领域的活动的蓬勃发展。

马拉尼马在定义上的判定，不断地创造着比我的分值要低的能量获取分数。而同时这种情况又不少见，那就是当他看到一些似是而非的猜测（例如，对于罗马帝国役畜的数量及因此而消耗的饲料量的人均数值，或者对人均消耗的木材量的猜测），而选择了较低的估计数字时，我们的差异还在进一步扩大。这些差异合起来，就会使我们每个人所估计的提高的幅度越发不同。

如果争论纯粹是定义上的，那倒不是很重要，因为马拉尼马和我都竭尽全力地做到了清楚、直白，读者们可以根据自己想了解的问题，选择使用相关指数。然而，马拉尼马还提出，我所得出的罗马帝国的数字肯定是夸大了。他认为我的数字意味着罗马人获得了比19世纪的许多欧洲人还要多的能量，而且罗马帝国早期的能量强度（即消费的能量与GDP的比值，或者基本上是耗费1千卡所能赚取的美元数）是1800年的西欧的两倍。

马拉尼马得出这样的结论，是将我对罗马帝国核心地带能量消费所做的每人每天31 000千卡估计，与他本人的欧洲人在1800年只能获取大约每人每天15 000千卡的估计进行了比较。他的数字比我得出的1800年左右西欧核心地带能量消费为每人每天38 000千卡的数字低了很多。我的数字来自库克、斯米尔和麦迪森的计算，因为纵观19世纪前的时期，马拉尼马定义的能量获取比库克的和我的都狭窄得多。结果，马拉尼马计算的1900年前的分数，始终只有我和库克计算的一半左右，所以，将我计算的100年的能量获取数与他自己计算的1800年的数字进行比较，只能产生荒谬的结果。我们的设想当然是不同的——按照马拉尼马的计算，西方核心地带的普通人在1800年消费的能量要比100年时的多大约75%；而依据我的计算，只多25%——但是，在罗

马帝国使用的能量比英国廉价的情况下，马拉尼马认为我的分数是荒谬的，纯粹是因为他坚持将术语定义不同的指数直接进行比较。

我们可以将定义上的分歧视为看待数据的两种不同方式，马拉尼马的定义将导致较低的结果，而我的定义则导致较高的结果。那么引人注目的是：总体的画面何其相似。马拉尼马和我都认为，欧洲从中世纪到工业革命之间的这段时期，能量使用是稳步增长的这种旧看法是错误的。能量获取状况自罗马帝国之后曾一路下跌，最快也要到1700年时，欧洲人才刚刚赶上罗马人的水平。甚至到20世纪时，我们的画面仍大致是相似的。依照我的定义，在西方核心地带，能量消费从1900年的每人每天92 000千卡增至2000年的每人每天230 000千卡（增长系数为2.5）；而依照马拉尼马的定义，在西欧，能量消费从1900年的每人每天41 500千卡增至2000年的每人每天100 000千卡（增长系数为2.4）。

然而，这种将马拉尼马的计算和我的计算相对比的方法，会忽略两者之间的巨大反差。正如我在前文中指出的，当我们只观察过去2 000年时，这两种方法会产生大致相同的情景。但是，当我们观察自最后一个冰期以来的历史时，情况就大为不同了。马拉尼马的数字意味着从拉斯科时期到泰斯塔西奥山时期，人均获取能量一定大致翻了一番（从养活能生存的人口所需的最低水平，每人每天4 000千卡增长至每人每天约8 500千卡），而我则认为应当增长了七八倍，增长至每人每天31 000千卡。

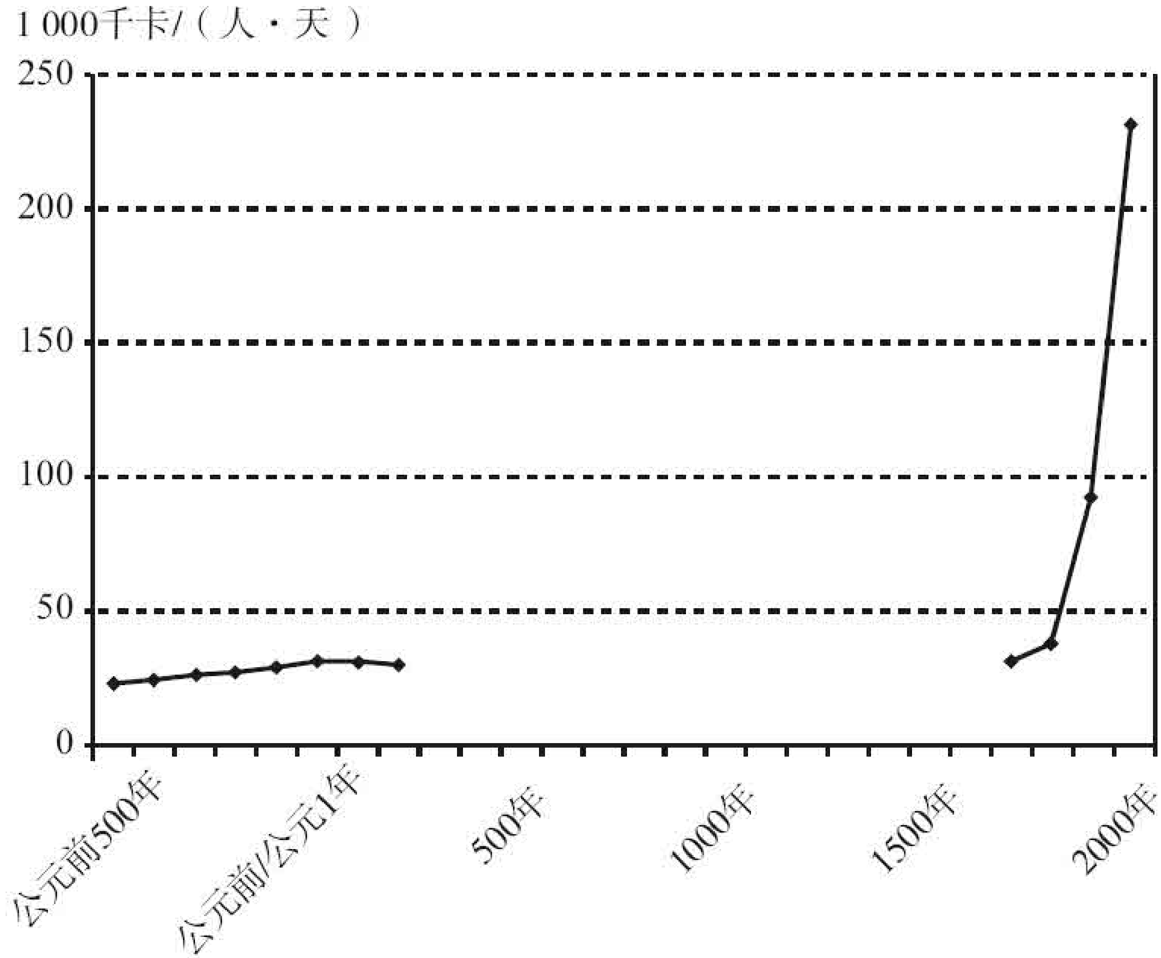
马拉尼马倒不像格里高利·克拉克那样离谱。正如我前面曾提到过的，克拉克曾说，“（公元）1800年时世界上的人均富裕状况并不比公元前100000年时的好。”然而，马拉尼马的数字意味着在公元前14个千年间，能量获取的年平均增长率只有0.005%。而我的数字意味着年平均增长率为0.02%，这个速度也不算快，但使前现代的经济发展展示出一幅非常不同且更真实的面貌。

在本章较前面的部分我曾提到过，对罗马世界感兴趣的经济学家们，通常都试图探究罗马的实际工资或人均GDP，却很少有人关心考古记录的杂乱的细节。其结果之一便是他们似乎往往对罗马世界与史前社会之间的鸿沟，缺乏比较清晰的认识。马拉尼马对度量什么和如何度量所做的设想，没有抓住罗马帝国的生活与像耶利哥那样一万年前的农业城镇的生活，以及像南非的平纳克尔角那样10万年前的定居点的生活之间的反差。观察非常长期的能量历史，需要更完备地研究考古记录，也需要采用像库克那样的先驱人物们创新的方法。能量流之类的问题对研究史前社会和古代社会至关重要，而库克等人的方法在解决这些问题时，比马拉尼马的方法要敏锐得多。

结论

贯穿公元前第一个千年，人均获取能量都有所增长，在1世纪时达到最高点，约为每人每天30 000千卡。按照前现代的标准，达到极高的水平，接近于1700年左右西方核心地带的水平，不过按现代的标准仍算很低，也许连当代美国15%的水平都达不到。图3.6显示了我对古代时期（公元前500~公元200年）和现代时期（1700~2000年）的估计。

图3.6 公元前500~公元200年和1700~2000年西方能量获取估计数字



古代和现代之间（200~1700年）

下一个挑战是填补古代地中海地区和早期现代欧洲的数据之间的漫长鸿沟。我把这1 500年分成了3个阶段：200~700年、700~1300年、1300~1700年。

衰退（200~700年）

图3.5显示了在这第一个阶段中，工业和商业活动出现了长达几个世纪的意味深长的衰退，说明能量获取也在下降。

原则上讲，罗马皇帝戴克里先于301年颁布的关于价格和工资的著名法令，应当能使我们了解4世纪初时的实际工资情况，从而为我们提

供起点，但实际上，问题远不是这样简单。根据沙伊德尔的计算，从该法令所推测的非熟练工人实际工资只相当于每人每天9 376千卡，低于1~2世纪时意大利的大约每人每天25 000千卡（不过存在很大的变量，12 000千卡上下）。然而，罗伯特·艾伦的计算显示，实际工资只相当于1 439千卡，和18世纪欧洲最贫穷的地区一样低，即使把工资全部用于购买食物，也维持不了太久。无疑，这个法令似乎说明从公元150~301年，实际工资下降了，但沙伊德尔和弗里森建议将其中的数字权当痴心妄想，与真实世界中的价格相距甚远，他们恐怕是对的。

最近几项对考古证据的研究，进一步加深了人们对公元200~700年间能量获取水平下降的印象，不过这些研究也表明，变化的详细情况和速度在地区与地区之间的表现殊为不同。一些新的能量获取手段，例如铧式犁和水车的使用，在公元200年后变得越发普遍，尤其是在罗马核心地带其他方面都相当落后的北部边缘地区，但是总体趋势是显著地在向另一个方向变化。

除非罗马晚期考古的专家们能够更精确地对考古证据进行量化，否则很难做出准确的估计，但是公元200~700年的概貌是：由石头和砖头盖成的高大房屋被木头和泥土建成的矮小房屋取代了；铺砌的街道被泥土道路取代了；排水沟和高架渠被废弃了；人的寿命、身高和数量都下降了，存活下来的人们从城市流向了农村；远程贸易衰落了；简陋的手制陶器取代了精致的轮制陶器；木制和骨制的工具使用增多，而金属工具减少；工厂纷纷歇业，被乡村匠人或家庭手工业者取代。

我在《西方将主宰多久》一书中曾提出，西方核心地带能量获取水平下降，是从1世纪60年代开始的。当时穿越大草原的人口迁徙，使得欧亚大陆东西端原本迥然不同的细菌融合了起来。图3.5显示，这种被称为“安东尼瘟疫”的突发疾病，在公元200年前就已经开始拉低能量获取水平了。随着气候开始恶化，到公元3世纪时，这种下降的趋势日

渐明朗，特别是在罗马帝国的西部；但是从5世纪开始的第二轮暴跌有着深远的影响。早在公元450年，不列颠岛的极西北部便显示出物质财富的骤降。到500年时，高卢也出现了这种情况。600年时蔓延到意大利和西班牙。到700年时又吞没了北非和爱琴海沿岸的拜占庭核心地带。

公元400~700年，崩溃的浪潮从西北席卷东南，通常呈现出各种纷繁复杂的模式，正如新近对6世纪安纳托利亚地区西部的萨迦拉索斯遗址的公共厕所进行的植物考古研究所证明的。研究显示在当地农业越来越精耕细作的同时，城市结构却在瓦解，呈现出明显矛盾的画面。然而，在长达三个世纪的过程中，整体效应却是明显无误的。西方的核心地带在地理上收缩了，缩小到今埃及、叙利亚和伊拉克一带。地域的缩小是与人均能量获取水平的下降是相应的。

这就是说，我们应当认为公元200~700年的西方核心地带能量获取水平的下跌并非是灾难性的。在埃及和伊拉克，灌溉系统、城市和基本的国家形态仍然是完整无缺的，阿拉伯人的征服也许反而刺激了农业生产的增长。在其他地方，即使在最黑暗的时代（如意大利地区的6世纪，或安纳托利亚地区的7世纪），人们仍然采集木头，生火做饭，做着与罗马帝国鼎盛时期大致相同的事情。然而，总体能量获取水平肯定是下降了。例如，最近对英国所做的稳定同位素分析表明，7世纪时，非常简单且千篇一律的谷物食物，代替了罗马时代更加丰富多彩的饮食。

在当前的证据状态下，我们只能根据特定的考古发掘报告，进行大致的猜想。从公元200~500年，能量获取水平可能下跌了10%（核心地带从大约每人每天31 000千卡降至大约28 000千卡）；然后在公元500~700年，又下跌了10%，降至大约每人每天25 000千卡。其实，公元200~700年，埃及和伊拉克的人均能量水平即使真有下降，恐怕也降得很少，但是意大利、北非和高卢南部等地区的暴跌，却导致了西方

最发达的核心地带的能量获取水平，在公元700年时比200年时低了20%。

这个下降程度远没有图3.5所显示的那样剧烈（原因是图3.5也许主要反映的是变化最大的非食物热量和昂贵食物热量），但在一些罗马史专家看来，似乎仍然被夸大了。在整个19~20世纪的相当长的时间内，历史学家们都倾向于认为爱德华·吉本（Edward Gibbon）对晚期古代社会的总体观点是正确的，但到了20世纪60年代，有批评家提出了反对意见。按照最主要的修正主义者彼得·布朗（Peter Brown）的说法："撰写晚期古代世界（公元200~700年）的历史简直是太容易了，好像那就是一个关于'衰落和瓦解'的悲伤故事。"布朗抛弃了吉本的黑暗画面，他主张："我们越来越多地了解到，这个时期是与令人震惊的新开端相连的……对于被一个敏感的欧洲人如此珍重地视为其自身文化中最'现代'和最有价值的……'当代'特性，我们表现得极度敏锐。"

布朗意在提醒历史学家们，不应让衰落和瓦解的叙事，遮掩了晚期古代社会文化转变的复杂而迷人的事实，但经过30年的提醒后，许多历史学家如今已走向了另一个极端。安德烈娅·贾尔迪纳（Andrea Giardina）曾评论说："现在有一种普遍的信念……认为诸如'衰落'或'颓废'之类的词意识形态色彩太过强烈，因而也就成了对人们的误导。"布朗主张我们应当将公元200~700年这个时期视为从古典文化进入早期中世纪文化的转型时期，这是非常正确的，但是太多历史学家任由这种新观念蒙住了自己的双眼，无视这也是政治和经济崩溃的时代这一事实。战略学家爱德华·鲁瓦克（Edward Luttwak）最近曾说，"新的时尚观点将晚期古代社会的转型几乎说成了和平的迁徙与温和的渐进式转变，然而却遭到了详细的考古证据的反驳。考古证据中充满了暴力、破坏，以及上千年都无法恢复的物质享受和教育成就的灾难性的损失。"我发现几乎没有人否认这个结论。

自20世纪60年代开始流行起来的对渐进主义模式的最好的修正办法，就是直接将对2世纪时罗马帝国任何部分的遗址考古报告和调查数据，与7世纪时同一地区的报告和数据进行对比。所有遗址（甚至包括比罗马帝国所有其他部分都更好地渡过危机的埃及的遗址）都显示出物质生活水平和能量获取水平的下降。

说不清的600年（700~1300年）

虽然很少有人质疑这600年间西方核心地带的能量获取水平是否在总体上呈现缓慢上涨的趋势，但详细情况却难以说清，很大程度上是因为研究中世纪伊斯兰世界的历史学家和考古学家，对于能量获取的关注度相比同行要低。

在公元700年时，西方核心地带已收缩至埃及－叙利亚－伊拉克地区。有一些证据表明，8世纪或9世纪是叙利亚，9世纪或10世纪是伊拉克，11世纪塞尔柱人入侵时是整个亚洲西南部地区，能量获取水平都下降了，但是在700~1300年这段时期，似乎埃及始终保持着能量获取的高水平，而西班牙地区有所上升。欧洲基督教地区在900年后显然出现了强劲的经济复苏，到1300年时，最富裕的地区意大利赶上了伊斯兰世界的核心地带埃及的水平。

拜占庭帝国在10世纪也显示出经济迅速复苏的迹象。经济学家布兰科·米拉诺维奇（Branko Milanovic）在其一篇极具价值的论文中，利用了相当丰富的资料，计算出在大约公元1000年时，拜占庭核心地带非熟练工人的平均实际工资为大约每年680美元（以1990年国际元的购买力平价）。像罗马的人均GDP的计算一样，这个数字对食物热量之外考虑很少。米拉诺维奇特别对非食物收入进行了小小的提高。然而，他的确注意到了，他得出的拜占庭人均GDP数字比大多数人估计的罗马帝国早期人均GDP数字低大约20%，比扬·范·岑登计算的1086年英国人的收入和格里高利·克拉克计算的13世纪初英国建筑匠的收入高20%~25%。所有这些人均GDP研究使用的都是同样的方法，说明即使

绝对数字低估了能量获取水平，随时间推移的相对变化仍然可以准确地反映实际状况。

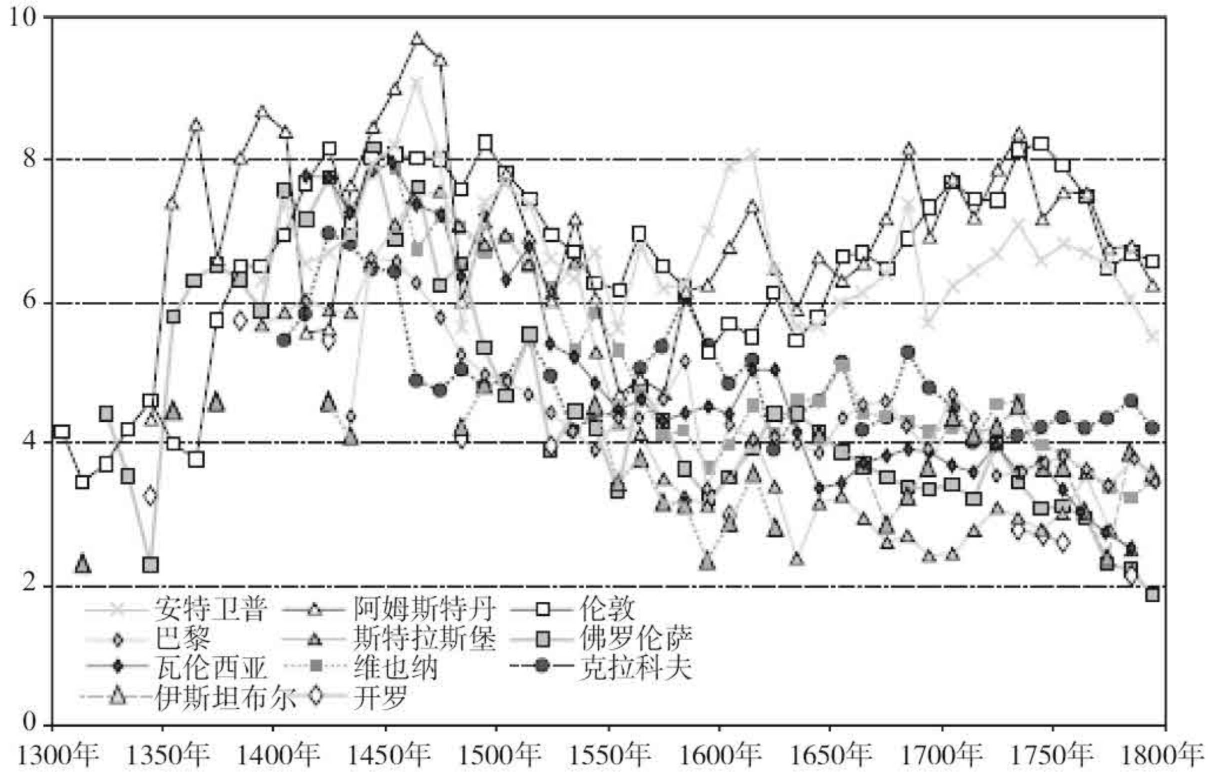
从这些对比推断，再为非食物热量做较大的提高，如果公元1世纪时罗马核心地带的能量获取水平大约为每人每天31 000千卡的话，那么公元1000年左右拜占庭的能量获取水平则约为每人每天26 000千卡。而如果米拉诺维奇所采纳的罗伯特·洛佩兹（Robert Lopez）的论断——公元1000年左右拜占庭和阿拉伯人的阿拔斯王朝的能量获取水平大致相当——是正确的，西方核心地带整体水平也应当达到每人每天26 000千卡。而公元第二个千年早期，英国遥远边缘地区的能量获取水平约为21 000千卡。如果有什么要补充的话，那么通过罗马和拜占庭人均GDP和实际工资的对比，也许会稍稍低估了公元100~1000年能量获取水平总体的下降情况，因为这一下降也许对非食物热量的影响比对食物热量的影响要大得多，而戈德史密斯–麦迪森–米拉诺维奇的估计，在很大程度上都忽略了这些非食物热量。

如果这一系列推断都是正确的，我们必然会得出结论：西方核心地带的能量获取水平在700~1000年只有极小提高，从每人每天25 000千卡增至26 000千卡。尽管这个数字无疑与希腊的发现是相符的，但考古证据的匮乏使得这点很难得到验证。核心地带的能量获取在700~900年比较平稳地保持在大约每人每天25 000千卡，在10世纪时开始攀升，在1000年时增至每人每天26 000千卡，然后可能在1300年时达到了每人每天27 000千卡。欧洲的考古证据似乎与此相符，有家庭财产增多、坚固房屋增多、贸易加强，以及政府开支大大增多的显著迹象。甚至在遥远的波兰边缘地区，11~12世纪的饮食也比以前要丰盛得多，花样也繁复得多，但是欧洲最富裕的地区似乎仍然是意大利。

我们不可能对13世纪的意大利和埃及进行直接的考古对比，这是令人沮丧的，但经济史学家谢夫凯特·帕慕克（Sevket Pamuk）搜集的实际工资数据显示，1300年时意大利北部的工资水平（可以推测总体

的能量获取水平）可能已经赶上了埃及，并领先于拜占庭；而到1400年时，意大利连埃及也超过了（见图3.7）。

图3.7 1300~1800年非熟练工人实际工资



资料来源：帕慕克，《黑死病》（*Black Death*），第297页，图2

快速增长（1300~1700年）

如果对中世纪和近代的能量获取水平的这些估计大致是正确的，那么在1300~1700年这段时期，西方核心地带的能量获取水平一定增长了大约23%，从大约26 000千卡增至大约32 000千卡。这比同样时间跨度的任何其他时期都要快，除了公元前400~前1年这段时间，当时增长了29%——从24 000千卡增至31 000千卡。古代和近代早期增长率和总体分数的近似，说明历史学家们喜欢将这些时期做类比，并非是不合时宜的。

欧洲许多城市自中世纪晚期以来的极其详细的实际工资系列数据，如今都可以得到。这些数据表明整个13世纪和14世纪早期非熟练

工人的工资总体减少了；1350年，黑死病加大了土地与劳动力的比率后，工资得到了巨大回升；随着15世纪晚期和16世纪的人口增长，实际工资又逐渐下降了，但是到1600年时，西北欧和南欧、东欧的工资水平拉开了距离：西北欧的工资呈回潮之势，而南欧和东欧仍在继续下跌；1700年时，阿姆斯特丹的非熟练工人的实际工资比在1350年时高出了30%，而伦敦的非熟练工人高出了80%，这两地的工资增长都高于前面提到的能量获取水平的提高。

安格斯·麦迪森对人均GDP的估计使得1500~1700年这段时期呈现出极其不同的面貌。按照麦迪森的计算，除了意大利，整个16世纪西欧的生产率全面持续增长。在他看来，1700年时荷兰和英国取得了领先地位，并非因为它们在17世纪复苏了而其他地区却倒退了，而是因为它们比欧洲的其他经济体发展得更快。他确定在1500~1700年西欧生产率增长了29%。

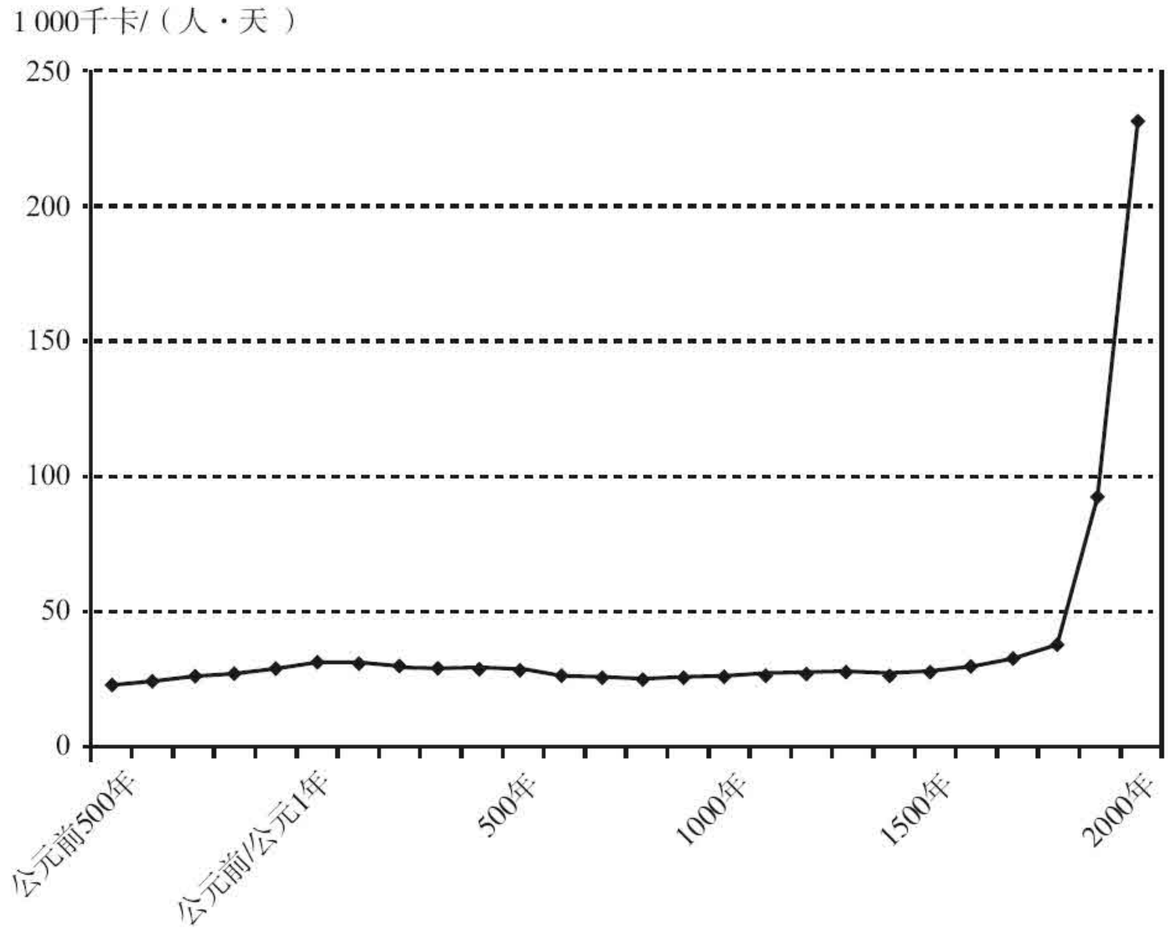
像古代时一样，实际工资和人均GDP情况之所以有所不同，很大程度上应解释为，度量的是非常不同的事物。西欧的领主们在黑死病肆虐后无法再任性妄为，使得资源配置大大地向穷人转移，实际工资的增长比生产率的增长快得多；但随着16世纪人口回升，实力又转回贵族那一边，尽管人均GDP仍在增长，实际工资却下降了。

1350年后长达一个世纪的实际工资飙升，也遮掩了14世纪的广泛大衰退的事实。衰退在很多方面损害了贸易和工业。20世纪90年代的研究表明情况并不像以往的某些历史学家认为的那样严重，但14世纪的灾难和不稳定似乎仍然拉低了能量获取水平，但下降幅度不大。从1300年的每人每天27 000千卡降至1400年的26 000千卡，但由于缺乏关于定居点的可量化的考古证据，这只是猜测。

1300~1700年的能量获取水平呈现上升（见图3.8），这方面的考古证据非常明显，似乎也符合我在前文中提出的23%的涨幅，尽管考古证据还没有详细到能够验证我所提出的每人每天的能量获取水平在

14世纪下降了1 000千卡的看法。西北欧农业产量增长的证据非常清楚，文字资料和实物证据都表明，13~14世纪在城市人口增长的压力下，捕鱼量有显著增长，渔场也得到了扩张。

图3.8 公元前500~公元2000年西方能量获取情况



食物热量的增加仍不足以明确地影响成年人的身高，但在非食物热量中，变化要惊人得多，特别是在1500年后。遗嘱和诉讼文书中显示的细节和发掘出的遗物，全都表明无论是在城市还是乡村，1700年时西欧的房屋比1300年时的更大、更精致，物质产品也更丰富。工业生产在增长，人们工作的时间更长，诸如泥炭和煤炭等化石燃料开始贡献巨大的能量。虽然对比肯定仍然是猜测性的，不会很精确，但17世纪时西北欧的人均能量获取水平可能已经超过了罗马巅峰时期（约公元100年时）。

冰期晚期的狩猎－采集社会（约公元前14000年）

尽管也许令人吃惊，但我们在估计冰期末期的能量获取水平时，较之其后到18世纪前的任何时代，我们有更充分的立足点。尽管自农民将最后的食物采集者逐出最早的西方核心地带（侧翼丘陵区）已过去了数千年，同时，该地区的气候和生态都发生了很大变化，比较研究仍能相当准确地确定能量获取的可能参数。

颇具盛誉的生物能量学和灵长类动物生态学等领域的研究，为我们展现了我们最近的进化邻居类人猿的能量使用情况的生动画面，而经济人类学家也测出了从炎热的非洲到寒冷的西伯利亚，当代世界各地的食物采集者的能量获取情况。

已知最早的作为物种的人，180万~250万年前生活在东非地区。他们的能量需求近似于黑猩猩，但有相当确凿的证据表明，这种“能人”吃肉的次数比黑猩猩要频繁，他们甚至可能已经成为活跃的猎手，而不是拾荒者。通常需要植物与太阳能的光合作用产生的大约10千卡的化学能量，才能使动物产生1千卡的动能，所以，能人已经在用昂贵热量取代廉价热量了。即便如此，能人由于身材矮小、脑容量较小、物质文化简单，可能通常也只需要大约每人每天1 500千卡的能量。

随着180万年前东非的能人进化成直立人，能量获取水平可能得到了巨大飞跃。直立人的脑容量增长了大约40%（从610毫升增至870毫升），体重增长了75%（从35千克增至62千克），身高增长了近50%（从1.15米增至1.7米）。直立人可能已能随心所欲地生火，这极大地增强了他们非食物能量获取的能力，并使他们通过改造食物从中吸收了更多热量。由于大约5万年前的考古记录太少，证据是存在争议的，但是最近在以色列的盖谢尔贝诺特雅各布地区的考古发现强而有力地证明了79万年前直立人已经懂得使用火了。假如从木头中释放能量烹饪食物已经成为尽人皆知的技术，那么直立人的总体能量获取水平可能已升至大约2 000千卡。

随着原始人类北迁越过北纬40°线，他们必须摄入更多能量以应对更为寒冷的气候。英国的山毛榉坑和德国的舍宁根都发现了4万年前的人类经常使用火的确凿证据。稳定同位素分析表明，尼安德特人摄入的食物能量中相当大一部分都是昂贵的肉类热量，特别是在较寒冷的地区和时间。生物能量学家估计他们通常每人每天要消费至少3 000千卡，也可能接近5 500千卡。

冰期晚期的现代人需要用于食物的热量相对较少，因此用于燃料的热量也较少，但其他类非食物能量的获取极大地增长了。对虱子的基因分析表明，人类至少在5万年前，也可能在15万年前，就开始穿合身的衣服了。对化石脚骨所做的解剖研究也表明，至少在4万年前，人类也普遍穿鞋了。大约5万年前，智人也开始将少量的能量用于个人装饰，将相对较多的能量用于建设栖身场所。考古学家们至今仍未发现关于原始人类盖房子的令人信服的证据，但至少自5万年前起，现代人开始将能量投入建筑了。从最早的时候起，修建这些建筑物就需要数千千卡的非食物能量，但是人们在没有山洞可以使用时，有了遮风避雨之所，就可以借助火炉取暖，使这种努力得到了回报。

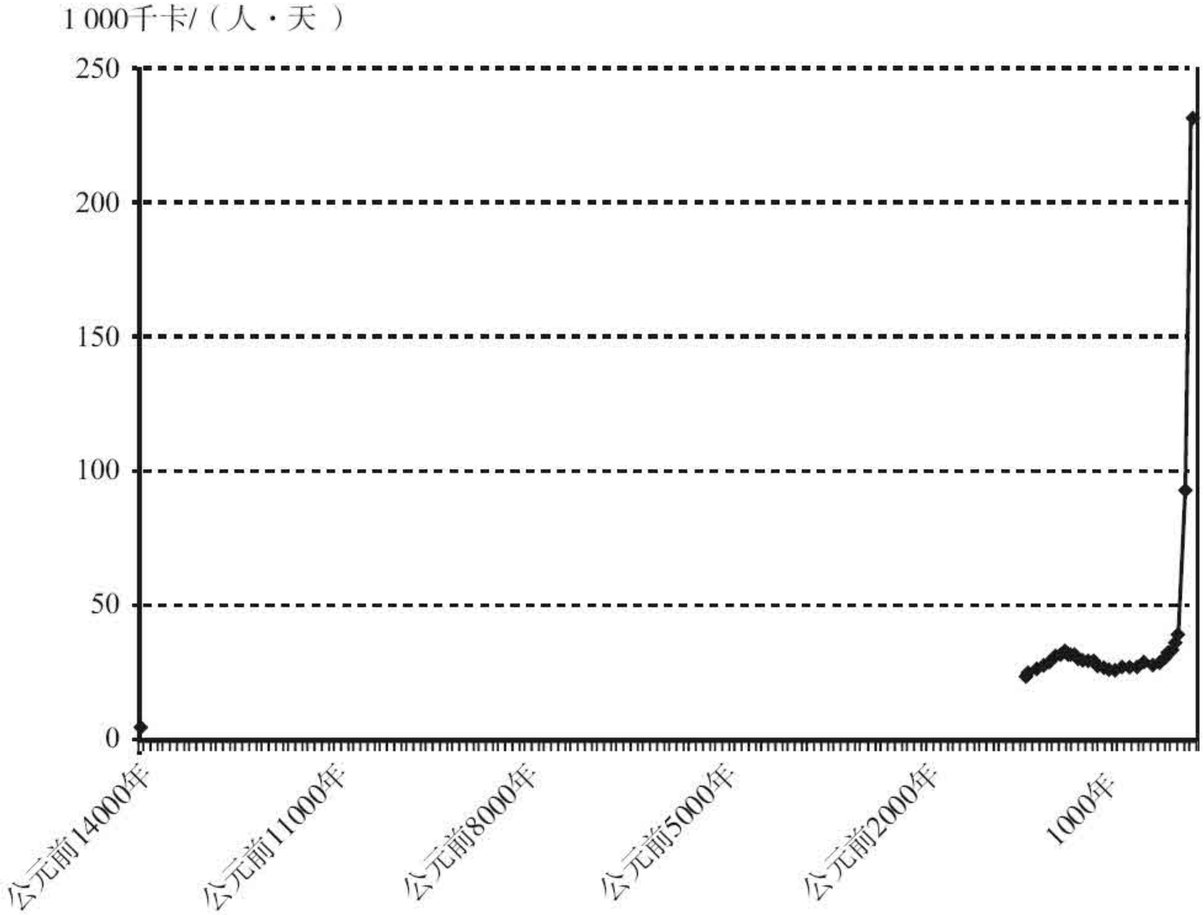
到冰期末期，大约公元前14000年时，在亚洲西南部西方核心地带，像欧哈罗这样的考古遗址，人类总体的能量获取水平（食物+非食物）一定是在每人每天4 000千卡左右。我做出这样的估计，是因为食物能量不可能长期远少于每人每天2 000千卡。假如非食物能量获取远低于此，纳吐夫人的物质文化就要比考古记录所显示的简陋许多；假如非食物能量获取远高于此，考古记录也会比实际显得富裕许多。

从采集者到帝国主义者（公元前14000~前500年）

正如图3.9所清楚地显示的，关于冰期晚期西方核心地带狩猎－采集社会能量获取水平合理有把握的估计数字（公元前14000年，每人每天4 000千卡），与下一个合理有把握的估计数字（公元前500年的地中海东岸城市居民，每人每天23 000千卡）之间，还有很大的空白需

要填补。我们可以为这13 500年简单地假设一个平稳的增长率，或者是算术的，或者是几何的。但实际上，将实地考古资料与文献数据相结合，借助经济人类学方法，再对比公元前500年之后的分数，我们是可以得到更准确的估计数字的（见图3.10）。

图3.9 公元前14000年和公元前500~公元2000年的西方能量获取情况



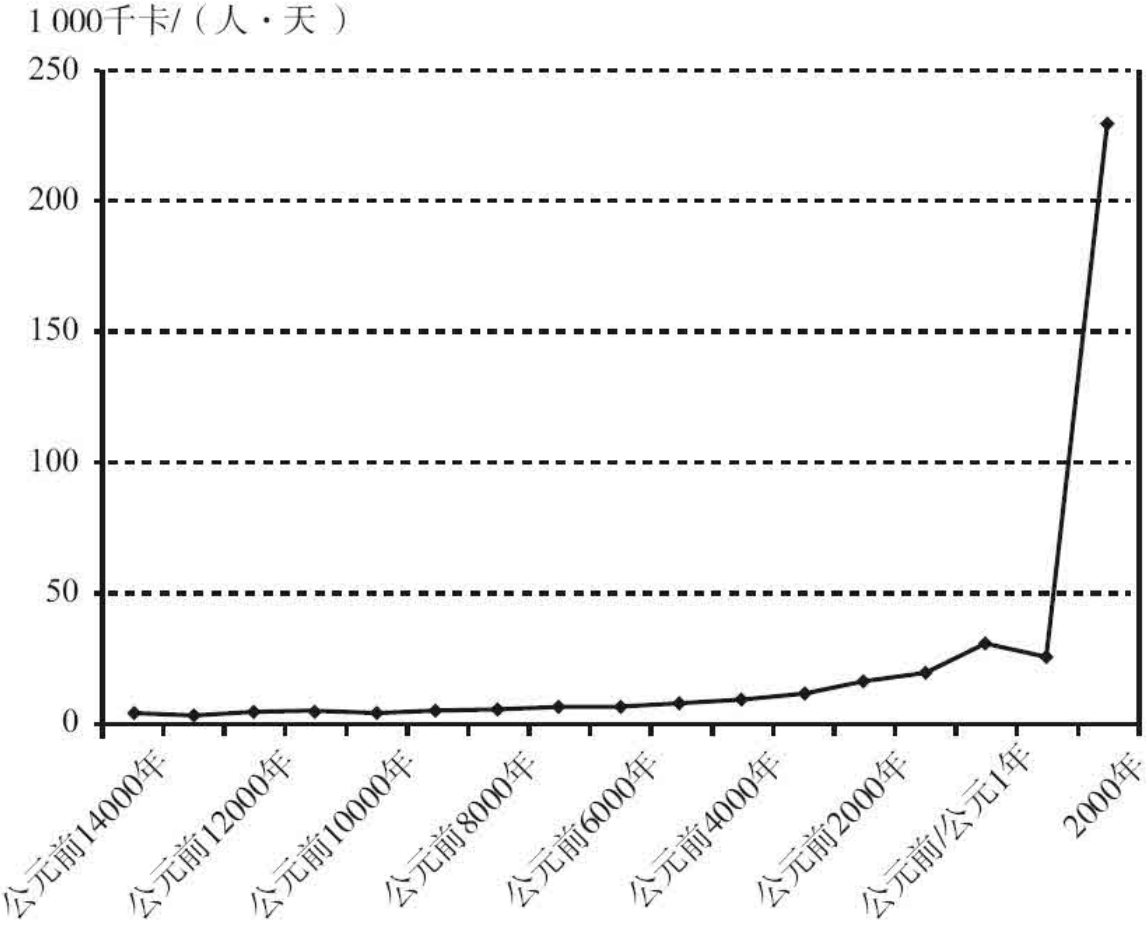
我将这一时期划分为6个阶段，先简要地概括一下每个阶段的大致发展情况，再努力对这些变化对于能量获取的意义进行定量分析。

丰富的采集社会（公元前14000~前10800年）

考古证据似乎非常清楚，亚洲西南部随着冰期末期气候越来越温暖、越来越稳定、饮食越来越丰富，棚屋越来越大也越来越精致了，物质文化在发展。叙利亚的阿布胡赖拉的考古发现表明，用于种植的

黑麦在公元前11000年时，更大种子已被选出来。亚洲西南部的人仍然是采集者（不过越来越多地成为定居者），到公元前11000年时，他们的能量获取水平仍然更接近于冰期晚期时的每人每天4 000千卡，而不是库克所认为的早期农业社会所应达到的每人每天12 000千卡，但是我们可以推测出，这3 000年至少就百分比而言有了相当大的相对增长，即便没有按照后世的标准，认为在千卡的绝对数量上有较大增长的话（见图3.10），也会得出这一推论。

图3.10 公元前14000~公元2000年每千年西方能量获取情况



新仙女木小冰河期（公元前10800~前9600年）

被称为新仙女木期的长达1 200年的小冰河期（公元前10800~前9600年），对能量获取意味着什么，是有争议的。一方面，公元前10000年时，很多永久性村落似乎被遗弃了，居民回到了更具流动性的

生活，在建筑和物质文化方面投入的能量也减少了；另一方面，最早的重大公共建筑出现在像柯尔梅斯德雷、杰夫–阿玛和穆赖拜特这样的考古遗址，意味着能量获取的增加。在我看来最稳妥的办法是，至少在相关证据得到重大完善之前，先假设能量获取水平在公元前10800~前9600年基本是持平的。这样与平稳的算术增长模式和几何增长模式都有了较大差异。这两种模式都认为公元前10800~前9600年之间能量获取水平增长了17%（算术模式认为从每人每天9 000千卡增至10 500千卡；几何模式认为从每人每天6 000千卡增至7 000千卡）。

农副产品的革命（公元前9600~前3500年）

公元前9600年后，随着天气变暖和定居生活的普遍化，我们看到两个截然不同的发展趋势。第一个是，耕种恢复得相当快。在公元前9000年时的约旦河谷、幼发拉底河谷和底格里斯河谷的多个考古遗址，都出现了非自然生长的小麦和大麦的大种子，到公元前8500年时就变得更普遍了，在很多考古遗址都能看到完全是人类种植的最早的小麦和大麦（坚硬的叶轴和外壳都还没有碾碎）。到公元前8000年时，沿着今伊朗、土耳其、叙利亚、黎巴嫩、以色列和约旦的边境地区延伸的侧翼丘陵地区，将近半数的碳化的谷物种子都是人类种植的了。到公元前7500年时，几乎全部的种子是种植的了。

驯化动植物的定居生活提高了平均每公顷耕地能量获取量，并且至少在短期内，提高了人均能量获取水平。然而，多余能量的主要用途之一是生养了更多婴儿，这就引发了第二个趋势。村庄陷入了“马尔萨斯陷阱”：几何级的人口增长超越了算术级的食物供给增长，使得人均食物供给跌回最低限度。这两个趋势共同造成了一种矛盾的结果，尽管公元前9600~前3500年非食物能量获取水平显然得到了巨大提高，总体食物供应的最好情况也就是停滞不前。种植谷物的廉价热量越来越多地替代了基于狩猎和采集野生食物的更为丰富的饮食。人们的骨

骼记录表明，总体而言，早期的农业人口健康程度不及农业产生之前的狩猎–采集群体。

过去30年的考古发掘情况也表明，自新仙女木事件后，能量获取的变化率比原先想象的要慢得多。我们不应想象发生了一场单一的农业革命，也许应当认为从用全部时间采集食物，到采集与耕种相结合，再到用驯化的植物和动物生产出的食物逐渐替代野生食物，是一个漫长的转变过程。最新的研究也表明，在侧翼丘陵区，这一过程大致发生于公元前9600~前7500年，经历了大约2 000年。

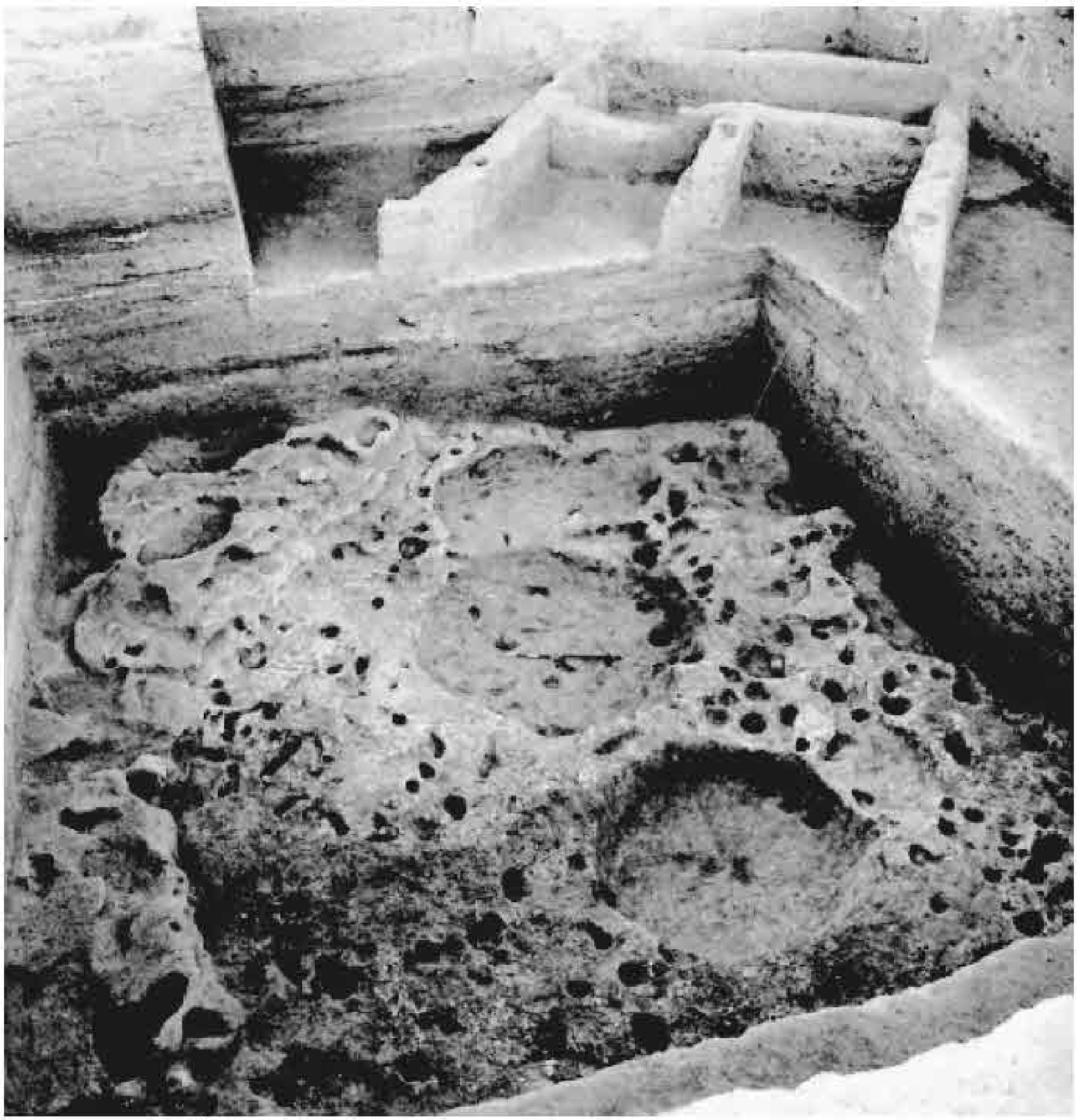
而且，这还只是第一阶段，驯化植物和动物这一转变之后，跟随而来的是时间更为漫长的食物能量的副产品革命。在这场革命中，农民们逐渐加强实践，发现了驯化的动植物的新用途。人们花费了很多世纪，才学会了轮换种植谷物和豆类以恢复地力；去除杂质，更有效地处理谷物；更有效地烘烤面包；约束动物用于取奶或牵拉，而不是在它们尚幼小时就把它们全部吃掉；制造更高效的犁和装有轮子的车；储藏设施也越来越完善，打井技术为溪水流不到的地方提供了水。

完整的古代干谷农业在亚洲西南部至少直到公元前4000年时才就位。到那时，除草、轮种和为庄稼施肥才全部成为常规做法，极大地增加了每公顷能量获取量，即使大部分甚至全部的能量盈余都转换成了更多的人口，而不是更多的人均获取食物能量。

非食物能量获取的增长依然缓慢，但已经明显得多了。就像计算后罗马时期的能量获取量一样，最好的办法就是对比不同时期的定居点遗址。来自阿布胡赖拉的一张非常著名的照片（见图3.11）很好地说明了这一点：顶部是一座小而结实的房子的一部分，建于公元前8000年左右；底部是一个脆弱得多的棚屋的残存，建于大约公元前12000年。如果我们将时间继续推移，我们发现在公元前6500年左右时，房屋更加牢固——最广为人知的例证是恰塔尔休于。到公元前

4500年时，美索不达米亚平原欧贝德时期越发令人难忘。迈克尔·罗亚夫（Michael Roaf）曾描绘过一个相当典型并且也保存得非常完好的例证：伊拉克的玛德胡遗址，面积170平方米。那时的房屋已经用泥砖建得很坚牢了，通常会围成一个阴凉的院子，有防水的屋顶，还有很大、很好的仓储设施。

图3.11 叙利亚阿布胡赖拉的房屋遗迹。底部为公元前12000年左右的棚屋的柱坑；顶部为公元前8000年左右的泥砖房屋遗存

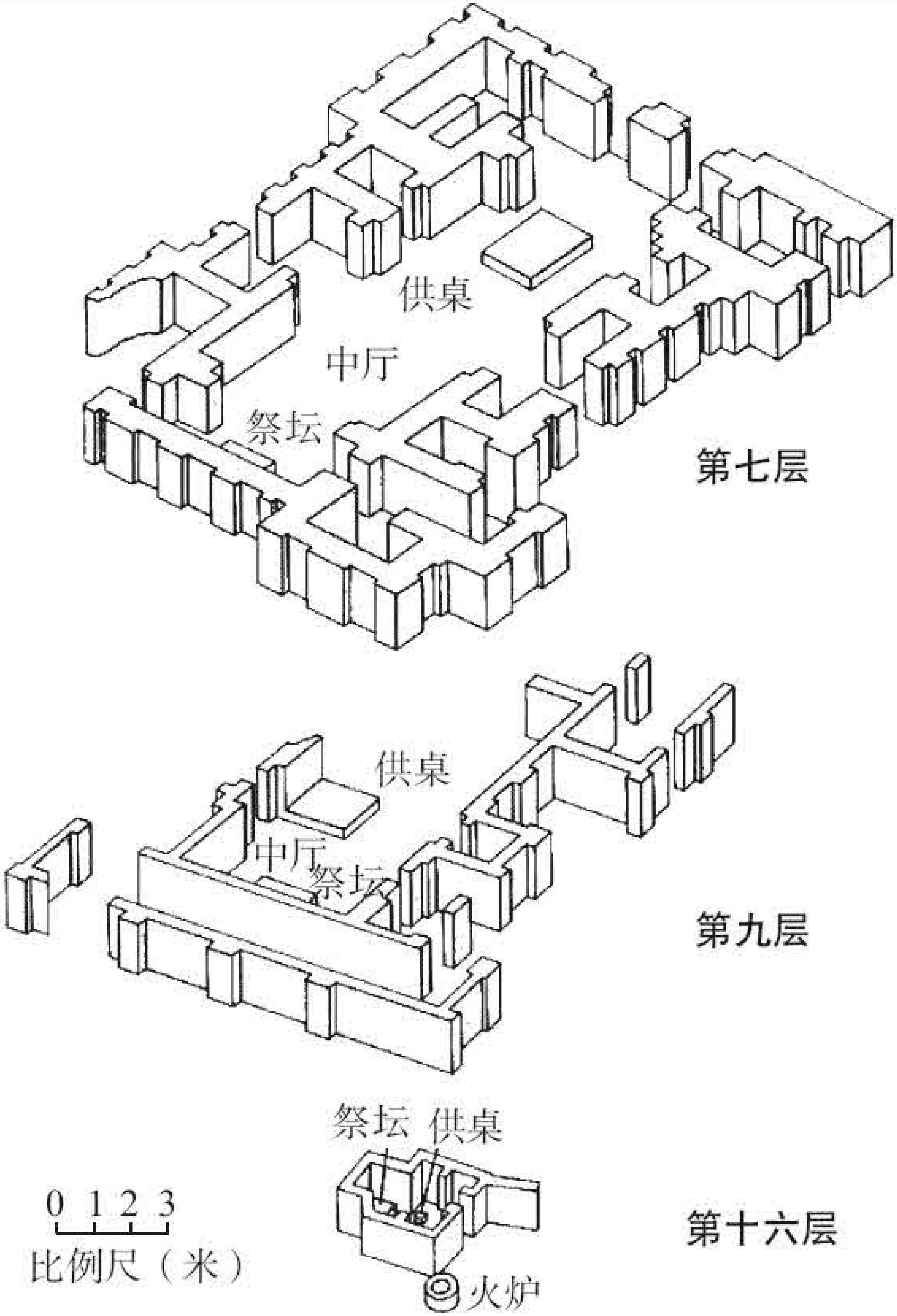


资料来源：摩尔（A.M.T. Moore）、莱格（A. J. Legge）和希尔曼（G. C. Hillman），《幼发拉底河畔的村落：阿布胡赖拉，从采集到农业》（*Village on the Euphrates：From Foraging to Farming at Abu Hureyra*）（2000年版），第107页，图5.4

典型的家居产品也同样在增长。陶器于公元前7000年左右开始使用，自那以后不久就有专业的制造者开始使用陶轮了。编织似乎也在稳步发展成熟，铜质的饰品、工具和武器于公元前3500年左右开始使用。据我所知，至今还没有考古学家对亚洲西南部不同历史时期的家居产品进行过系统的定量分析和比较，但阿布胡赖拉（约公元前12000~前8000年）和玛德胡遗址（约公元前4500年）家居产品内容的反差，还是很引人注目的。

各种各样的重大公共建筑所消耗的能量也急剧增长。耶利哥有一些早至公元前9000年的作为防御工事的塔，但是与公元前3500年埃利都的精美神庙或者苏萨的堆积的巨大土台相比，就要逊色多了。图3.12是约公元前5000~前3500年埃利都神庙序列的复原图，能够像阿布胡赖拉的房屋图片一样形象地说明非食物能量获取的增长情况。

图3.12 伊拉克埃利都神庙遗迹。最底部是建于约公元前5000年的神庙；最顶部是建于约公元前3000年的神庙


供桌
中厅
祭坛
第七层
供桌
中厅
祭坛
第九层
祭坛
供桌
火炉
第十六层
0 1 2 3
比例尺（米）

资料来源：迈克尔·罗亚夫，《古代美索不达米亚和近东文化地图集》（*Cultural Atlas of Ancient Mesopotamia and the Near East*），1990年，第52页

用于运输的能量也在增加。将畜力与带轮的车辆联系在一起的最早的确凿证据，是在公元前4000年左右苏美尔人所绘的牛车图。而在公元前3000年左右的坟墓中，出土了真实的车辆。风力和水力也得到了利用。独木舟于公元前5000年被用于打鱼；在埃利都，公元前4000年时就开始使用正经的船了。

公元前9600~前3500年非食物能量获取水平增长非常明显。不过，我们应当记住，就公元前14000~前11000年丰富的采集社会而言，尽管公元前9600~前3500年能量获取水平在百分比上变化很大，但在绝对数字上，按照现代的标准却并不大。甚至到了这段漫长的历史时期末期，西方核心地带的人们仍然是村民，他们的能量获取水平大致处于图3.1中库克所谓的“早期农业社会”阶段。

古代国家（公元前3500~前1200年）

公元前3500年后，随着有中央政府的国家的创建和扩张，西方核心地带的能量获取增长率开始加速。但由于没有对人体骨骼资料进行系统性搜集，讨论又一次无法进行，就像对旧石器时代饮食的稳定同位素分析及其他人类学分析不足一样。零散资料给人们形成的总体印象是，食物热量摄入的总体变化相对不大。

对于西方核心地带不同部分的饮食和营养，我们可以勾勒出非常简略的画面，但细致研究后会发现，地区间的差异非常大。高产可能是长期趋势：公元前第三个千年和第二个千年之间的结实率上涨了（在公元前2000年时美索不达米亚的灌溉大麦种植可能达到了30:1），但人口似乎增长得同样快，抵消了收益。

然而，像较早的时期一样，人均非食物热量获取也有极大增长。最引人注目的是金属器具的普及。这决定了这个时代的标准名称：青

铜时代。据皇家官方档案记载，宫廷里有很多青铜铸造坊；考古发掘者也发现了大量的私人铸造坊；自公元前1200年后，西方核心地带的石制工具在很大程度上就消失了。

著名的埃及金字塔、亚述和巴比伦金字形神塔，以及青铜时代的宫殿和神庙，无疑消耗了巨额的能量。约公元前2600年修建于埃及吉萨的大金字塔，至今仍然是世界上最重的建筑物，重约100万吨。长距离贸易的规模也在迅猛扩大，尤其是在公元前1600年后，在土耳其海岸外发现的船只残骸可以清楚地证明这一点。不过，最重要的还是公元前第三个千年和第二个千年大量新增人口所消耗能量的增长。公元前3500~前1200年，在西方核心地带的所有部分，房屋标准和家居产品的数量和工艺都有较大的改善。

像其他时期一样，这段时期内也存在强烈的地区差异，以及突发的地方性崩溃事件。例如，爱琴海克里特岛的新宫殿时期（约公元前1800~前1600年）是一个十分富足的时期，有很大的房子（中等规模房屋的楼层平面都有130平方米）和丰富的物质文化。然而，公元前1600年后，克里特岛上的非食物财富似乎下跌了，而与此同时希腊大陆却仍在持续增长。

这一时期最大的突发性崩溃事件，似乎发生在公元前3100年后的美索不达米亚。乌鲁克被焚毁了，其庞大的物质文化区瓦解了；还有公元前2200~前2000年，发生在从美索不达米亚平原经叙利亚和黎凡特到埃及的整个地区（地中海沿岸很多地区也受到了冲击）。然而，尽管这两次崩溃都留下了清晰的考古痕迹，但却没有明显证据表明能量获取受到了很大影响。

这似乎有好几个原因。其中一个很大的原因是两次崩溃实际上发生的地点都非常分散，一些遗址被毁灭并放弃了，而另一些却依然繁荣（例如在叙利亚，今塞赫那城和斯维哈特大约在公元前2200年被放

弃了，而布拉克丘和莫扎却发展得更大了）。对于崩溃的根本原因，考古学家们意见不一，有些人甚至质疑使用“崩溃”这个词是否合适。

另一个原因是公元前3100年时在埃及出现了新的核心地带。尼罗河谷没有受到公元前3100年崩溃的影响，尽管公元前2200年后的灾难的确给埃及造成了重大冲击，但与对美索不达米亚平原的作用时间不同。公元前2100年，埃及的古王国和美索不达米亚的阿卡德王国都已经瓦解了，但强大的新国家乌尔第三王朝重新统一了美索不达米亚大部。到公元前2000年时，乌尔王国也崩溃了，但中王国重新统一了埃及。尽管公元前2200~前2000年这个时期留下了明显创伤，但西方核心地带的能量获取水平似乎始终在增长。公元前1800~前1550年又一轮动乱发生时，能量获取似乎仍在增长。

还有一个原因，我界定能量获取的办法也许低估了危机对此的影响。在早期历史的这段时期，公元前2500年前，我每500年计算一次分数；公元前2500~前1500年之间，我每250年计算一次分数。公元前3500年的分数度量的是乌鲁克崩溃之前的能量获取水平。虽然美索不达米亚的能量获取水平在公元前3000年时也许低于公元前3100年之前（证据并不清楚），但埃及的能量获取水平在公元前3000年时肯定比美索不达米亚在公元前3500年时高。同样，对公元前2250年的能量获取的计算，显示的是大崩溃开始前的水平，而且虽然美索不达米亚在公元前2000年时仍处于混乱中，埃及的秩序已经恢复了。正如我在第二章中提到的，在两个分离甚远的时间点之间进行度量，不可避免的结果就是抹平现实的变化。

既然我们为城市化设置了低门槛（大约5 000人），可以说到公元前1250年时，西方核心地带的很多人都已经成为城市居民了，而欧亚大陆西部的大多数人已经生活在有行使职责的中央政府的古代国家中了。他们的能量获取水平已经远远超过了库克的图表（见图3.1）中早期农业阶段的水平——每人每天12 000千卡了。即使将最富裕的青铜

时代晚期定居点，如乌加里特（毁灭于约公元前1200年），与一流的希腊定居点，如奥林索斯（毁灭于公元前348年）相比，青铜时代社会的能量获取水平仍无法与古希腊的每人每天大约25 000千卡相提并论。

青铜时代末期（公元前1200~前1000年）

公元前1200~前1000年，席卷了整个西方核心地带的崩溃提供了能量获取水平下降的第一个无可争议的证据。在受危害最严重的地区（今希腊和土耳其），城市和精美的经典建筑一起灰飞烟灭，甚至在受危害最小的地区（埃及），社会上层的活动也锐减了。

至今尚无很多证据证明埃及普通百姓的生活有什么变化，但在叙利亚、以色列和爱琴海地区，住房的标准、物质产品的数量和质量，以及贸易网络的规模都急剧下降了。缺乏人体骨骼的大规模系统性对比，又一次造成了问题，但至少在爱琴海地区，成年人的死亡年龄下降了；同时，发病率增长的一些证据也肯定了成人身材呈下降趋势。

铁器时代早期（公元前1000~前500年）

能量获取水平一定是急剧飙升了，才从公元前1000年左右青铜时代后的低谷期蹿升至公元前500年的约25 000千卡。而公元前500年，正是地中海古代社会经典时期的开端。

大多数可用的资料和早期阶段所用的同属一类。像通常一样，经典建筑是最明显的证据：公元前6世纪波斯波利斯的波斯王宫和巴比伦的神庙、宫殿，使此前几个世纪的所有建筑都黯然失色，就像以弗所的阿耳忒弥斯神庙和罗马卡皮托林山的朱庇特神庙对扩张的核心地带边缘地区所起的作用一样。

在核心地带，房屋说明不了太多问题。那里有多间屋子、方方正正的房子一般面积都在50~100平方米。这种情况好几个世纪来都很普遍，但在以色列，结实的两层楼“柱屋”在公元前1000~前500年变得更普遍、更大，也更豪华。再往西到地中海地区，多间屋子的方正房子逐渐取代了面积较小、曲曲弯弯、只有一间屋子的房子。这一过程在希腊始于公元前750年左右，到公元前500年时大致完成；在意大利南部和西西里，始于公元前600年左右，到公元前400年时告一段落；在法国南部始于公元前400年左右，到公元前200年时基本完成。

在希腊，关于成人身材的证据多少有些模糊不清，但公元前1000~前500年，成年人的平均寿命明显提高了，发病率也可能下降了，说明作为其基础的能量获取水平也在增长。然而，关于饮食的直接证据仍然不清楚，因为目前遗址中在食物方面的变量使得仍不可能探求现象变化。

还有一个非常引人注目的变化是铁的普及极大地提高了肌肉力量的效用。在公元前第二个千年很早的时候，金属就已经偶尔被使用了，但公元前1100年后不久，塞浦路斯地区的锻工们开始更加系统地打造铁器。这可能是因为公元前1200年贸易道路崩溃后，获得制作青铜器所需要的锡出现了困难，不得不采取对策。到公元前800年，贸易在很大程度上恢复了之后，铁的优势（尤其是其含量丰富和成本低廉）已经尽人皆知了，于是铁被继续当作工具和武器的标准原料。公元前1000年，希腊几乎所有的武器都是铁制的了；公元前700年，最早的铁制工具也在希腊出现了。当时铁制武器在意大利、法国南部和西班牙东部也已得到普遍使用。

公元前800~前500年，对地中海中部和西部地区的兼并，是西方核心地带最迅速的扩张。尽管亚洲西南部的老核心地带的经济活动无疑仍在增长，但希腊、意大利、西班牙、法国南部和今天突尼斯所在的

地区发展得更为迅猛。最容易量化的证据是船只残骸和污染记录（见图3.5）。

对希腊以外地区的人体骨骼、住房以及其他形式的证据缺乏系统性的搜集，又一次妨碍了估算，但总体面貌似乎是清晰的：公元前1000~前500年，西方核心地带的能量获取水平在上升——也许像以前一样快。在地中海盆地的中部和西部，上升得尤其快。

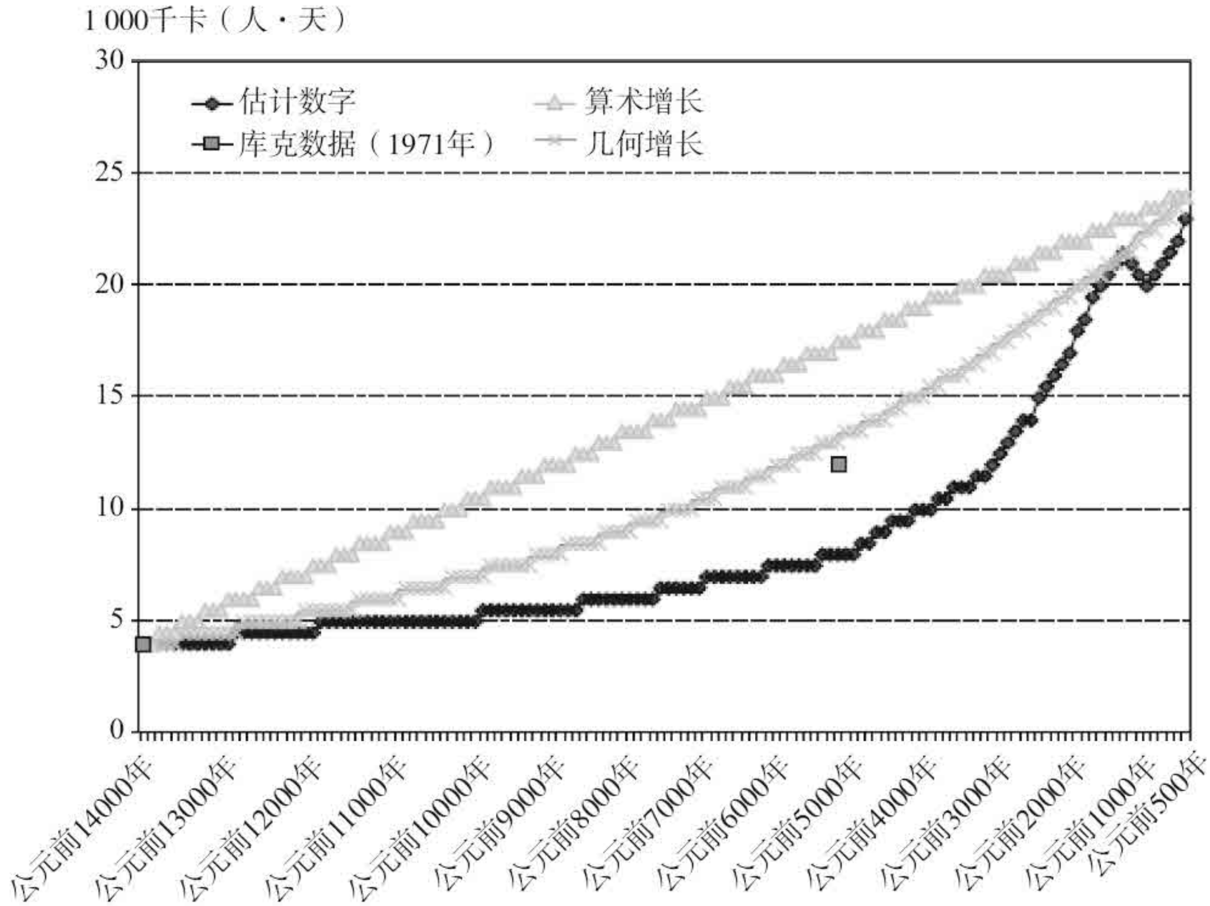
能量记分牌

要填补公元前14000年（每人每天4 000千卡）到公元前500年（每人每天23 000千卡）之间13 500年的能量获取分数的鸿沟，办法之一是索性假设一个固定的增长率，无论是算术的还是几何的（见图3.13）。然而，本节讨论的证据显示，那样做将遗漏大量重要信息。

考古证据显示得非常清楚，公元前最后几千年的能量获取水平，比冰期晚期和冰期刚刚结束的时期，增长得要快得多，意味着算术增长曲线肯定存在严重的误导。如果设定一个几何级的固定增长率（每年0.013%），会更大致符合实际，但即便如此，也会遗漏一些重要的细节，如公元前10800~前9600年新仙女木期能量获取的中断、约公元前3500年后的显著加速，以及公元前1200年后的下跌。最好的估计曲线无疑是几何曲线和算术曲线下方的那条曲线。其增长率是指数的，但指数总体是随时间推移而增长的。

然而除了这些基本的看法，我们没有确凿的一个点，继续推进的唯一办法是估计，然后将估计结果与实际的考古证据、比较证据，以及我们为公元前500~公元2000年期间估计出的分数进行对比。

图3.13 估计公元前14000~前500年西方能量获取水平的可供选择的办法



公元前14000~前10800年，能量获取水平也增长了，但极其缓慢。像新仙女木期前的阿布胡赖拉这样的定居点，相比于像欧哈罗这样的冰期晚期的考古遗址，人们获取的能量似乎更多。我猜测增长大致在每人每天1 000千卡，从每人每天4 000千卡增至5 000千卡（即3 200年间增长了25%，平均每年增长0.007%）。我这样假设并没有确凿的根据。房屋面积和精致程度、食物准备的复杂程度和物质文化的扩展程度的增长，也许只代表每人每天500千卡的增长（即12.5%的增长率）；也许能代表每人每天2 000千卡的增长（即50%的增长率）。虽然这两个数字都过于极端了，但即使两者中有一个比我估计的每人每天1 000千卡更接近于现实，公元前14000~前10800年的变化量仍然非常小，假设公元前10800年时的能量获取水平为每人每天4 500千卡，或者每人每天6 000千卡，而不是每人每天5 000千卡，对于随后的计算的影响都微乎其微。

正如前面的章节提到的，新仙女木期（公元前10800~前9600年）的证据有相互矛盾的迹象，所以我决定将这1 200年的能量获取情况视为没有变化。当然，这也许又是错误的。能量获取水平也许下降了（不过不可能一直跌到公元前14000年的水平），也许在继续上升（不过不会像公元前14000~前10800年那样快）。不过，像较早的时期一样，所涉及的数量微乎其微，估计误差更可能是相互抵消，而不大可能相互掺和。

公元前9600~前3500年，能量获取水平的增长似乎远大于公元前14000~前10800年。库克估计能量获取水平在公元前5000年时已经升至每人每天12 000千卡了，只比上面的几何曲线所指示的每人每天13 000千卡低一点。但目前可用的考古证据表明，这似乎太高了。库克也许犯了20世纪中叶的考古学家们偶尔会犯的错误，以为农业革命只是一场单一的、极其迅速的转型。然而我们现在已经知道，耕种和驯化是一个长达大约4 000年的过程，而且这还只是在亚洲西南部一直持续到公元前4000年左右的不间断的革命副产品的第一阶段。在公元前9600~前3500年，总体的能量获取水平大致翻了一番，从大约每人每天5 500千卡增长至每人每天11 000千卡（年增长率为0.013%，几乎比公元前14000~前10800年翻了一番），而不是像库克认为的那样到公元前5000年时就翻了一番还要多。按照库克的估计，公元前10800~前5000年的年增长率为0.017%。如果以这样的年增长率，到了公元前3500年，分数将达到每人每天15 500千卡。按照我的设想，如果能量获取在公元前3500~前1200年几乎又翻了一番，那么青铜时代晚期的能量获取水平就将达到每人每天30 000千卡——几乎相当于1世纪罗马帝国全盛时期、12世纪中国宋朝，或者大约1600年时西欧和中国核心地带的水平了。

这恐怕根本不可能。假如库克估计的公元前5000年的为每人每天12 000千卡的数字是正确的，那么与后来的数字保持看似合理的关系的唯一办法，就是假设公元前5000年后能量获取的增长率急剧减缓。

假如增长率降到只有每年0.015%（也就是说，比库克估计的公元前9600~前5000年的0.017%的年增长率还要低），那就将把公元前12000年的分数拉低至每人每天21 000千卡，像我估计的一样。然而，考古证据却很难与公元前5000年后增速减缓的说法相吻合。库克对大约公元前5000年西方核心地带能量获取的估计（每人每天12 000千卡）一定是过高了。如果能量获取在公元前10800~前5000年大致增加了50%，也就是从每人每天5 500千卡增加到大约8 000千卡（而不是像库克认为的那样翻了一番还要多，从每人每天5 500千卡增至12 000千卡），然后在公元前5000~前3500年又增长了大约1/3（从每人每天8 000千卡增至11 000千卡），那么我们绘出的新石器时代的能量使用情况，及其与青铜时代的关系的图画，就合理多了。所以，公元前5000年时，能量获取水平增长至每人每天8 000千卡，然后到了公元前3500年，增长至每人每天11 000千卡。

公元前3500~前1300年——大致是从乌鲁克时代到拉美西斯大帝的时代——能量获取水平大约又翻了一番，从每人每天11 000千卡增长到21 500千卡（年增长率为0.029%，恰好是公元前9600~前3500年增速的两倍，是公元前14000~前10800年增速的4倍）。如果这个判断是正确的，我所估计的增长曲线就在公元前13世纪赶上了几何曲线（见图3.13）。公元前1300年的数字当然有可能或高或低，但任何较大的改变（比如说低于每人每天18 000千卡，或高于每人每天25 000千卡）都将意味着在公元前第一个千年的早期，必须假设一个奇慢或奇快的变化率。

公元前1300~前1000年，能量获取水平的下降程度是比较难以估计的。我曾经提出，公元前13世纪的这个数字稍稍下降了一点，从每人每天21 500千卡降到21 000千卡，然后下降速度加快，在公元前1200~前1000年，从每人每天21 000千卡降到20 000千卡（公元前1200~前1000年的年下降率为0.025%）。波谷的底部也许还要更深一些，那样的话，公元前第一个千年早期的增长一定要稍微快一些，以便在公元

前500年时达到每人每天23 000千卡；波谷的底部也可能稍微浅一些，那样的话，随后的增长一定会更慢一些。然而，20世纪90年代，一些考古学家声称公元前1200年后崩溃的规模实际上极小，甚至根本没有崩溃，是误入歧途了，好比提出后罗马时代也没有崩溃一样。

如果这些数字大致是正确的，那么能量获取水平在公元前1000~前500年，一定是上升了大约15%，从每人每天将近20 000千卡增至23 000千卡（年增长率为0.029%，比公元前3500~前1200年的估计增速稍快一些）。据我的估计，能量获取水平在公元前500~前1年又上升了35%（从每人每天23 000千卡增至31 000千卡）。

在图3.5中，船只残骸和铅污染分别代表着远程贸易和金属制造业。公元前第一个千年中，15%的增长出现在公元前500年之前，另外85%的增长在公元前500年之后完成。这也许意味着我对公元前1000年（其实是公元前1300年）的估计太低了；或者这也可能只是反映了一个事实，公元前第一个千年的地中海地区出现大规模的人口增长（沙伊德尔估计在公元前1200~公元150年，人口大致增长了4倍），其中大部分是在公元前500年以后增加的，意味着贸易和工业的总量增长似乎在公元前第一个千年晚期得到了很大的重视，但人均增长量却没有得到很大重视。

惊人的积累：西方能量获取

图3.2和图3.3显示了我为公元前14000~公元2000年整个时期西方能量获取水平所计算的分数。这样的图表就其本质而言，大量数据都是近似值。很难想象每个数字都是正确的，这意味着（如第二章中所提到的）应当问的问题不是所有的数字是否都正确——我们可以肯定地说不是这样——而是这些数字是否错误到严重歪曲了西方能量获取的历史轮廓的地步。

对于这个问题，我想答案一定是：否。我计算的分数无疑都是在正确范围内的，出于我在《西方将主宰多久》一书中讨论过的原因，系统性误差的幅度恐怕不会超过上下20%。然而，最受关注的肯定是非系统性误差在多大程度上扭曲了图表的形状。

图3.14 公元前14000~公元2000年西方能量获取情况（假设罗马的增长率较低而早期现代时期的增长率较高）

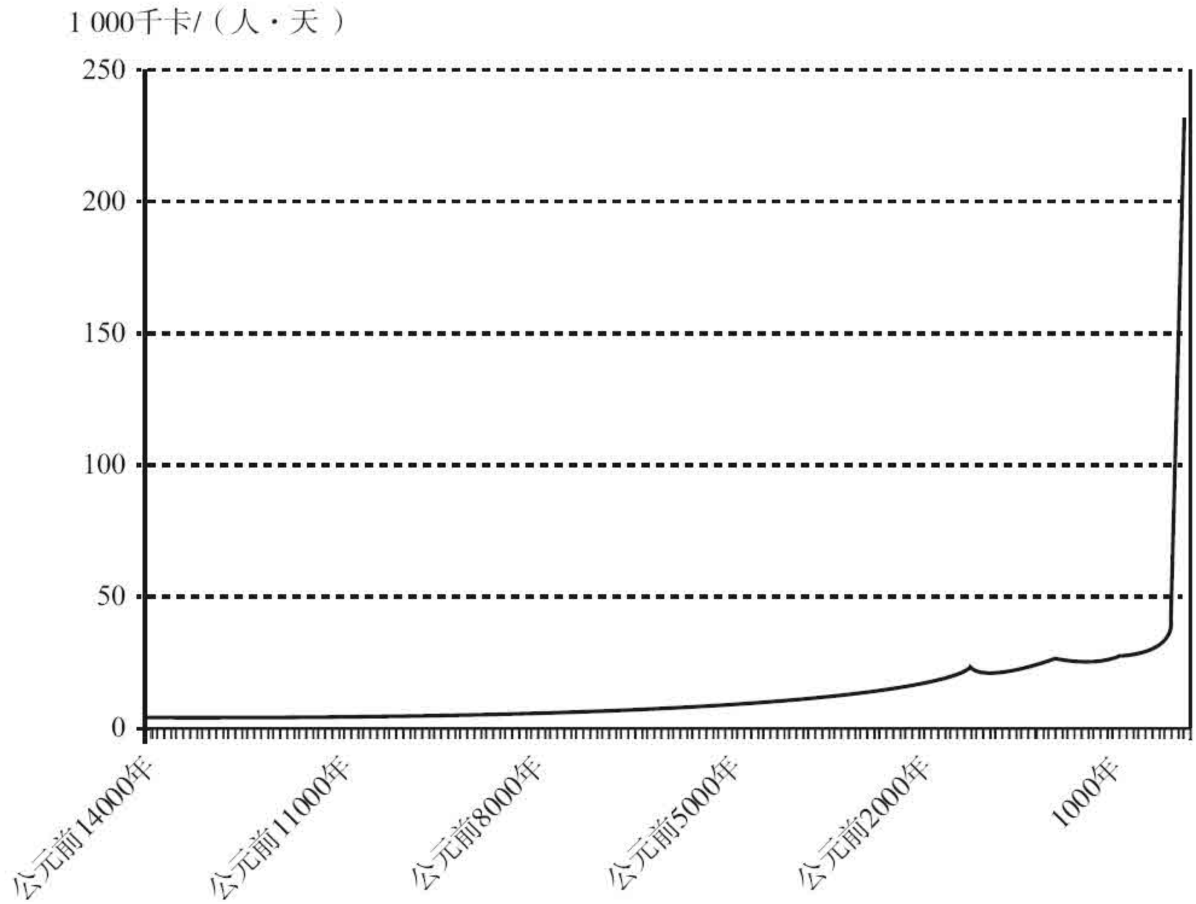
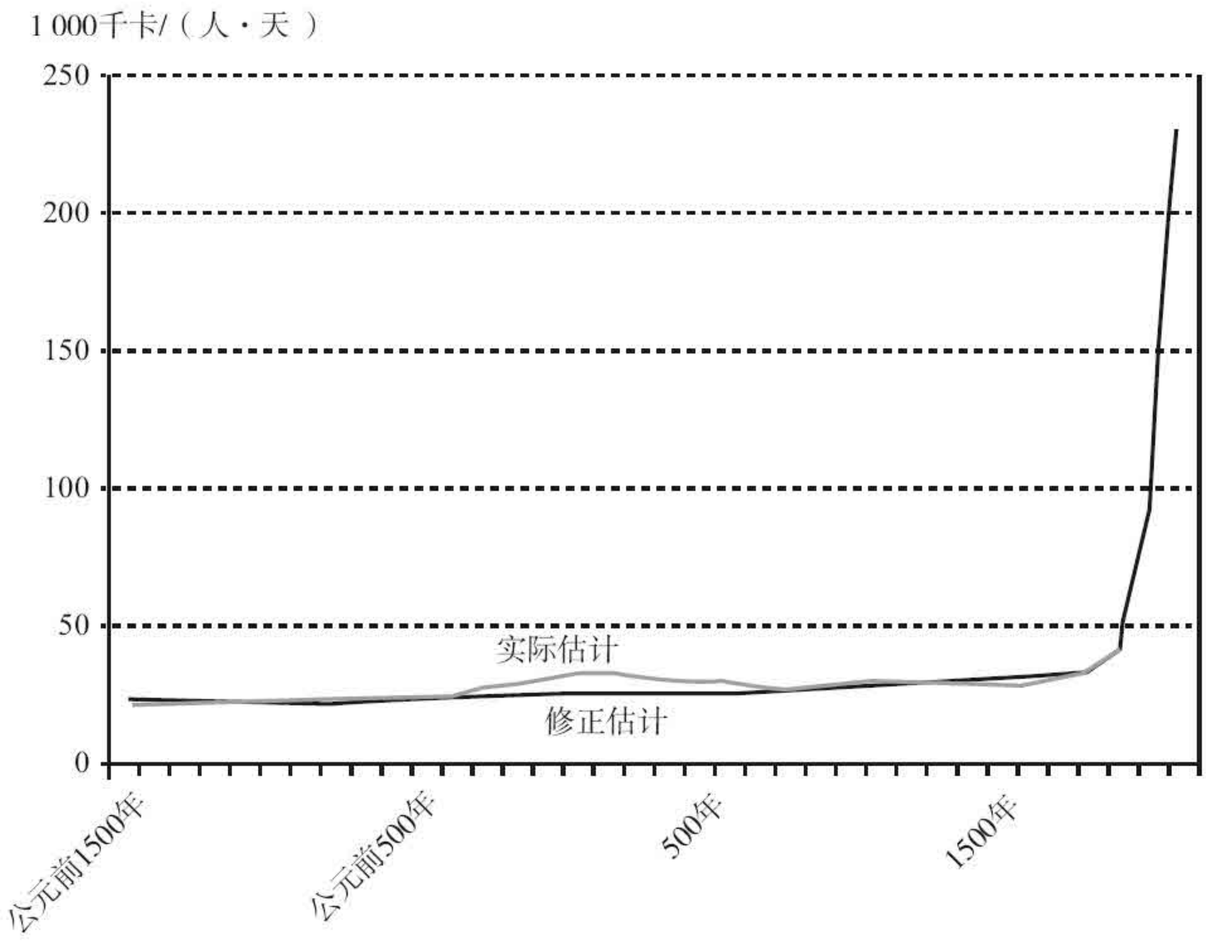


图3.14显示了能量曲线将会如何变化，例如，公元前第一个千年的能量获取水平实际上只是我估计的一半（即在公元前1000~前1年，从每人每天20 000千卡增至25 500千卡，而不是31 000千卡），而公元700~1500年的增长，则是我估计的两倍（即从每人每天25 000千卡增至29 000千卡，而不是从每人每天25 000千卡增至27 000千卡）。这些重大的修正令我非常吃惊，因为难以得到现存的考古证据的确证；然而它们对图3.14的影响却并不大。公元前1000~公元2000年的能量获取

增长情况变得平稳多了（这从图3.15中更容易看出。该图既显示了实际估计，也显示了这些修正估计，而且只显示公元前1500~公元2000年这一时期），但基本情形大致是相同的。

图3.15 对公元前1500~公元2000年西方能量获取情况实际估计与修正估计（假设罗马时期的分数应降低而早期现代的分数应提高）的对比

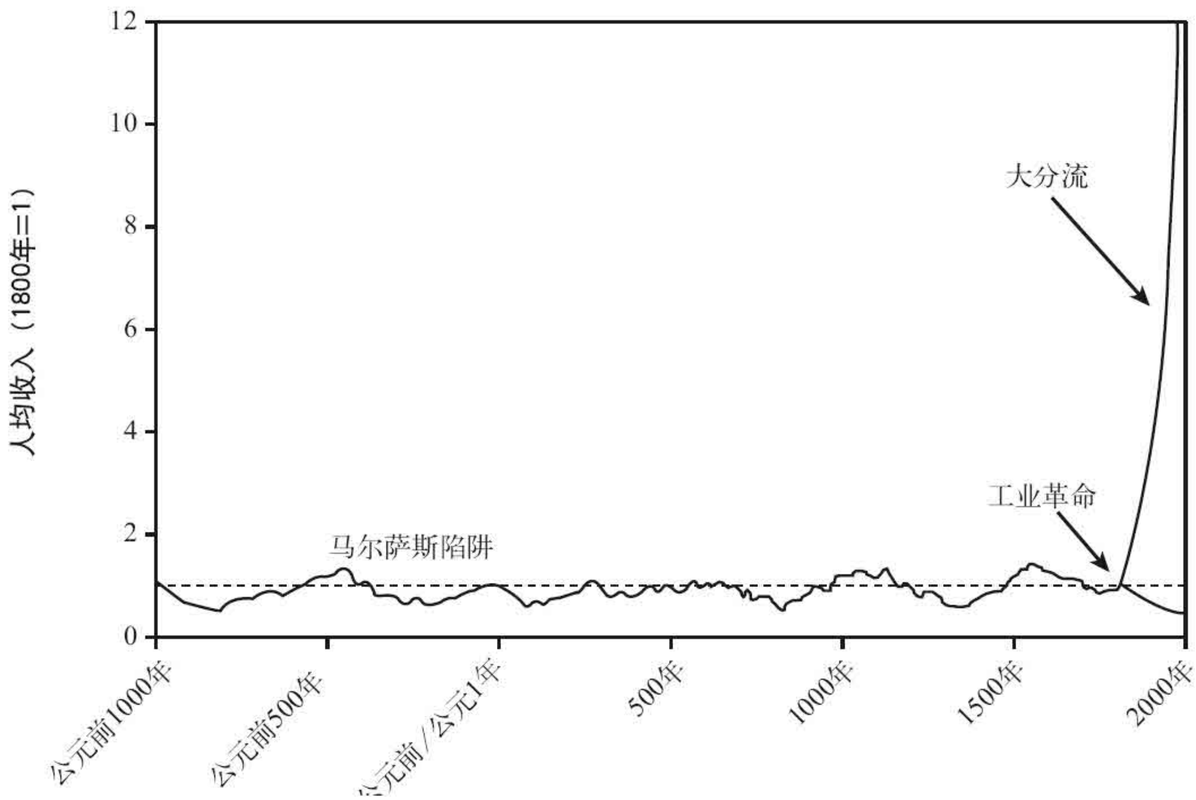


我们可以拿任何假设的修正数字来做实验，但是这样的实验的主要价值就是展现我们需要对分数进行多么剧烈的改变，才能对西方能量获取历史的基本形态产生重大影响。基本的模式是非常清楚的——从冰期末期到国家开始出现（从约公元前14000~约前3000年），是一段极其漫长的发展极度缓慢的时期，在早期国家和帝国时代（约公元前3000~前1年）虽有所加速，但发展仍然非常缓慢，然后波动不断冲击着农业社会的上限，使之稍稍超过了每人每天30 000千卡（约为公

元1~1600年），继而是短暂的农业社会的上限继续被推动提高的时期（1600~1800年），最后是（至今仍在持续的）爆炸式增长的短暂时期（1800年至今）。

经济学家们经常认为工业革命前没有发生任何重要变化。格里高利·克拉克曾说过（本章前面曾引用过）：“（公元）1800年时世界上的人均富裕状况并不比公元前100000年时的好。”他配发的图表（见图3.16）将前现代时期的生活标准表现为在马尔萨斯论的上限周围盲目乱撞，其非同凡响之处只在于其直率，观点本身仍然是错误的。

图3.16 格里高利·克拉克对过去3 000年来人均收入的再现



资料来源：格里高利·克拉克，《应该读点经济史》（A *Farewell to Alms*），第2页，图1

从冰期末期到1800年，能量获取水平有巨大的增长。然而，正如马尔萨斯本人所认识到的（见第三章第一小节），能量必须区分为食物热量和非食物热量。当人们将能量的意外收获转换为生养更多的孩

子时，每单位土地上食物热量的增长很快就被消费掉了。然而非食物能量获取却不会被抵消掉，考古记录证明，过去16 000年的积累是惊人的。图3.2和图3.3中的上升趋势被各种各样的崩溃打断，其中最引人注目的是公元前1200年后、公元200年后和公元1300年后，但这样的中断每次都被证明是暂时性的，并且只能毁灭先前增长的部分成果。

后来者居上：东方能量获取

对东方能量获取情况的研究，比对西方的要少得多，定量估计尤其少。然而尽管还有许多工作需要做，东方能量获取情况的主体轮廓却相当清晰。在最后一个冰期末期，大约公元前14000年时，在东方那些最得天独厚的地区，人均能量获取水平与西方是不相上下的，都在4 000千卡左右。出于地理原因（就此而言，生态差异使欧亚大陆西部可驯化的潜在动植物物种比东部要多），东方的分数起初上涨得比西方慢，耕种和驯养植物最早的清晰迹象的出现，比西方核心地带晚了2 000年。东方的分数增长到公元前3000年时开始加速。像西方一样，在公元前第一个千年的早期发生了一次严重崩溃。但东方的能量获取水平迅速恢复了，到公元400年时又开始向上提升了，但直到公元1000年后，达到农业社会的上限——大约每人每天30 000千卡。在公元1200~1400年间又经历了一次严重崩溃之后，东方的分数于1600年又恢复到农业社会的上限，然后在18~20世纪迅速增长（相对于更早的时期）。表3.4、图3.17和图3.18显示了我对自公元前14000年起东方能量获取情况的估计。

相对而言，在史前社会和古代社会，以及19世纪和20世纪，东方核心地带的分数似乎比西方核心地带低，但在西方历史学家们称之为中世纪和早期现代的时代，也就是大致从公元第一个千年的中叶到第二个千年的中叶，东方核心地带的分数要高。不过，将这些比较再进行精练，要困难得多。

在本节中，我首先从最近的时期——1800年开始。接下去，像我在分析西方能量获取情况时一样，我先从广为人知的时代（先是960~1279年的宋朝，然后是公元前206~公元220年的汉朝）开始，然后再填补鸿沟，最后再转入史前时代的东方。

表3.4 公元前14000~公元2000年东方能量获取情况

年代	每人每天能量（千卡）	分数（分）	年代	每人每天能量（千卡）	分数（分）
公元前14000年	4 000	4.36	公元前500年	21 000	22.88
公元前13000年	4 000	4.36	公元前400年	22 000	23.97
公元前12000年	4 000	4.36	公元前300年	22 500	24.52
公元前11000年	4 000	4.36	公元前200年	24 000	26.15
公元前10000年	4 000	4.36	公元前100年	25 500	27.79
公元前9000年	4 500	4.90	公元前/公元1年	27 000	29.42
公元前8000年	5 000	5.45	公元100年	27 000	29.42
公元前7000年	5 500	5.99	公元200年	26 000	28.33
公元前6000年	6 000	6.54	公元300年	26 000	28.33
公元前5000年	6 500	7.08	公元400年	26 000	28.33
公元前4000年	7 000	7.63	公元500年	26 000	28.33
公元前3500年	7 500	8.17	公元600年	27 000	29.42
公元前3000年	8 000	8.72	公元700年	27 000	29.42
公元前2500年	9 500	10.35	公元800年	28 000	30.51

公元前2250年	10 500	11.44	公元900年	29 000	31.06
公元前2000年	11 000	11.99	公元1000年	29 500	32.15
公元前1750年	13 000	14.17	公元1100年	30 000	32.69
公元前1500年	15 000	16.35	公元1200年	30 500	33.24
公元前1400年	15 500	16.89	公元1300年	30 000	32.69
公元前1300年	16 000	17.44	公元1400年	29 000	31.06
公元前1200年	16 000	17.44	公元1500年	30 000	32.69
公元前1100年	16 500	17.98	公元1600年	31 000	33.78
公元前1000年	17 000	18.52	公元1700年	33 000	35.96
公元前900年	17 500	19.07	公元1800年	36 000	39.23
公元前800年	18 000	19.61	公元1900年	49 000	53.40
公元前700年	18 500	20.16	公元2000年	104 000	113.33
公元前600年	20 000	21.79			

图3.17 公元前14000~公元2000年东方能量获取情况（线性–线性标尺）

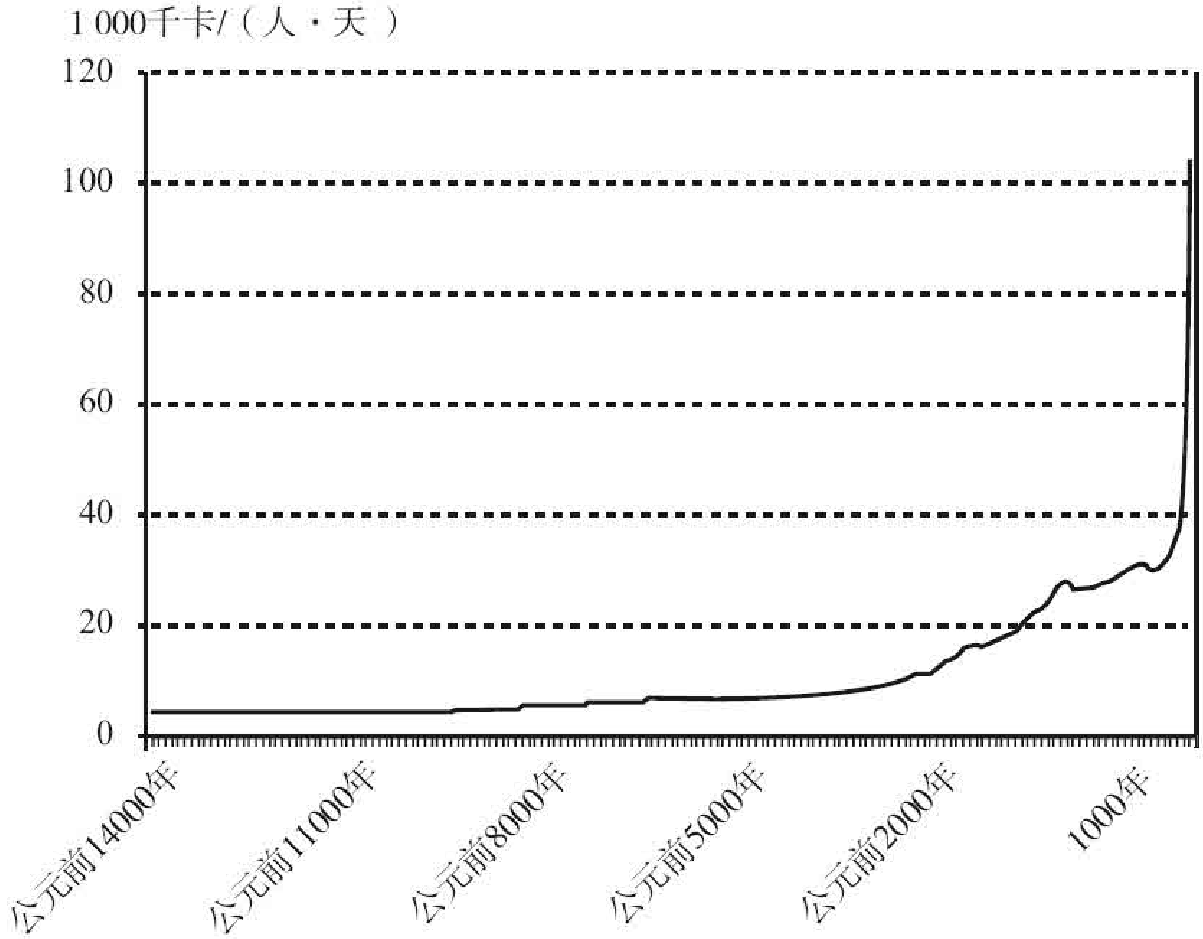
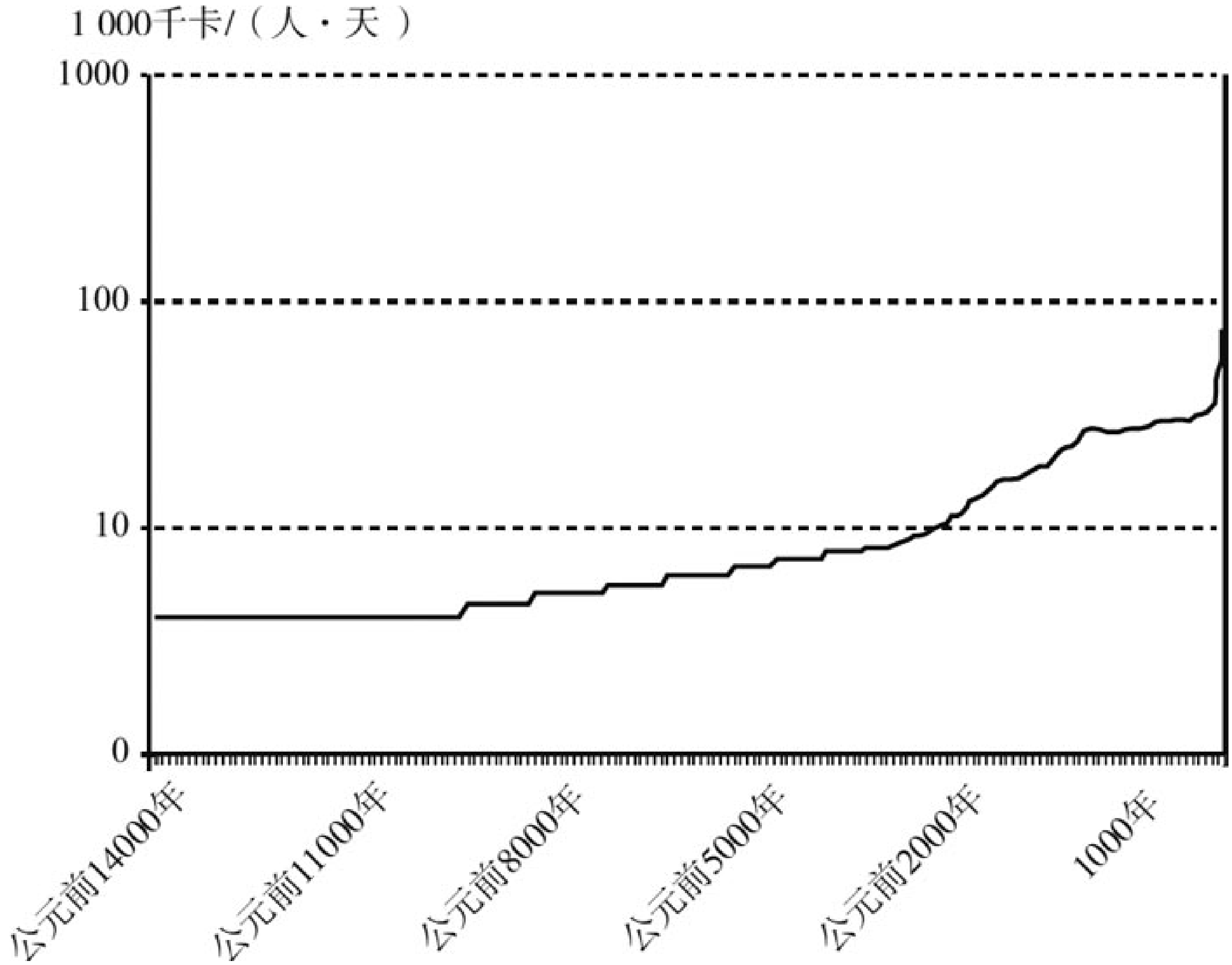


图3.18 公元前14000~公元2000年东方能量获取情况（对数–线性标尺）



最近的过去（1800~2000年）

像西方核心地带一样，2000年时东方核心地带也有高质量的统计资料，可供计算能量获取情况时使用。东方核心地带（日本）的人均食物热量和非食物热量的总体能量获取水平为大约每人每天104 000千卡，不到美国的消费量每人每天230 000千卡的一半，但比东方（或西方）历史上任何较早时期高得多。

在东方，可靠的政府统计数据不会在历史上回溯太远，而且（像西方一样）农民家庭用于燃料、房屋、衣服等的生物量缺乏量化的资料，使得问题更为复杂。1900年，日本烧了300万吨煤（大致相当于每人每年消费500千克，或者略高于每人每天500千卡；而1903年英国的消费水平为1.81亿吨煤＝每人每天4.36吨≈每人每天40 000千卡），以

及只能称之为微量的石油。然而，随着18世纪和19世纪人口压力的增大，生物燃料的使用也越来越高效和密集，而资源基础在逐渐弱化，也许堪与十八九世纪西北欧先进的有机经济体相比。总之，各种不同的资料表明，1900年东方核心地带日本的能量获取水平低于每人每天50 000千卡。

20世纪初期中国北方农民的生活相对记录得比较好。煤和豆渣肥料在19世纪时已经得到广泛应用。1900年时的生活水平普遍低于日本，一些地方实际上还在下降，但能量获取水平肯定高于每人每天40 000千卡。

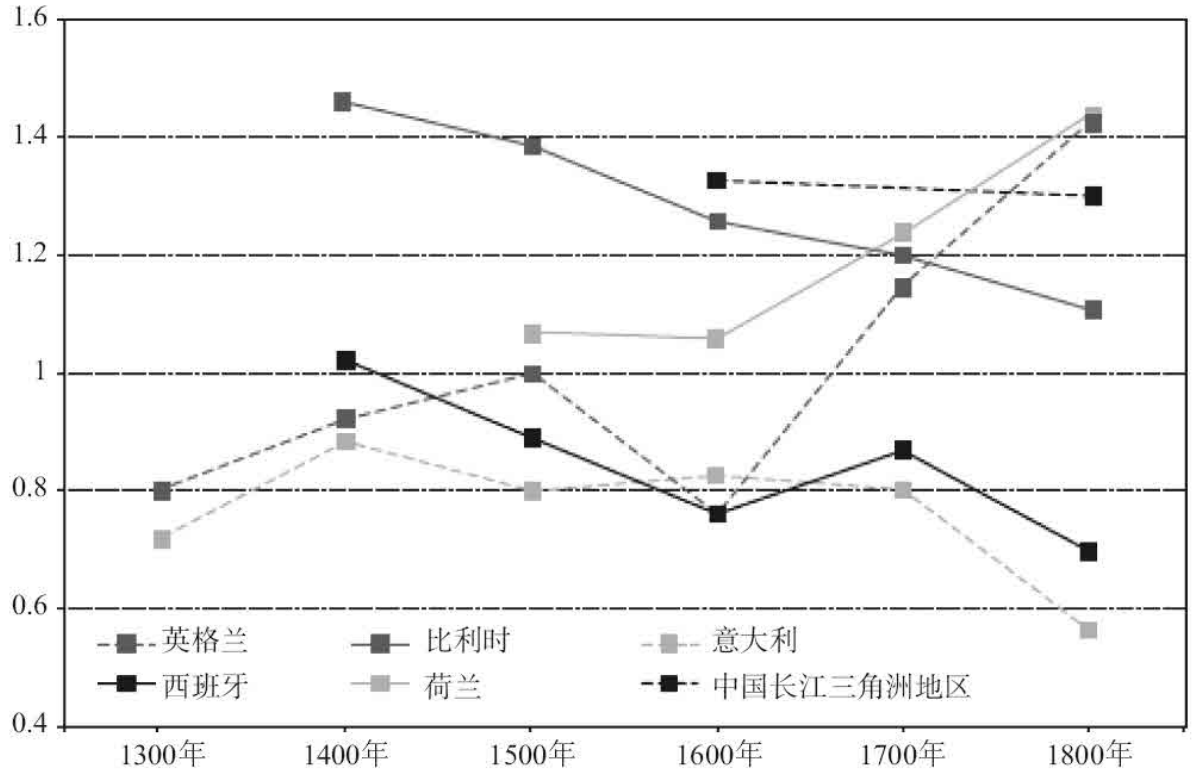
19世纪东方（尤其是中国）的生活水平，自20世纪90年代以来一直存在激烈争议，以致成为西方统治的长期注定论和短期偶然论争执的主要战场。20世纪的大部分时间里，历史学家中主导的理论是，中国经济在1400~1900年是停滞的。例如，安格斯·麦迪森估计，中国的人均GDP在1000~1500年，从450美元增长到了600美元（以1990年国际元表示的购买力平价计算），然后在1500~1820年，一直停留在600美元没有变化。同样，德怀特·珀金斯（Dwight Perkins）认为，中国农业在中世纪经历了强劲的增长和变革后，于元朝（1271~1368年）达到了其极限，自那以后，最优的实践方法从中国南方农业核心地带传遍了全中国，但基本上没有再出现重要的新技术。伊懋可（Mark Elvin）提出了一个更宽泛的观点，中国在宋朝（960~1279年）接近于工业起飞后，就陷入了一个他所谓的“高水平均衡陷阱”，传统的依靠肌肉力量和水力的技术达到了其效率的极限，但却没有足够的诱因跃升到使用化石燃料的技术。无论是含蓄还是直率，此类观点都认为东方核心地带的人均能量获取水平，从1368年明朝建立到19世纪40年代欧洲人入侵，基本上没有什么变化。

这些理论在20世纪90年代遭到了猛烈攻击，部分是因为中国于20世纪80年代向学者们开放了大量明朝档案，历史学家们从中发现了中

国经济发展的丰富证据，特别是清朝（1616~1911年）的证据。彭慕兰尤其主张，18~19世纪中国经济最发达的地区（长江三角洲）的发展轨道，与西欧的发展轨道更多的是相似而不是差异。他认为，中国原始工业化形式与工业革命是类似的。彭慕兰还提出，清朝时期的中国即使人口迅速增长，生活水平还是在提高中。按照他的计算，19世纪典型的中国成年男子每天消费2 386~2 651卡路里的食物，与英国的水平是大致相当的。中国人消费的食糖、烟草、蜡烛、家具和肉类的数量似乎也在上升中，棉质衣服已普及寻常百姓。

为1400~1900年的东方勾勒出的农业退化的悲观的旧画面，仍然有其捍卫者，但随着东方实际工资和农业产量的长期资料的完善，越来越多学者认为，在两种理论中折中是最为合理的。悲观主义者们认为，1600~1800年农业劳动者的人均产量的确是下降了（见图3.19），但仍然相当高，晚至1700年，长江三角洲地区的农业劳动者的生产效率可能仍然比欧洲任何地区的都要高。

图3.19 1300~1800年欧洲和中国的农业生产力情况



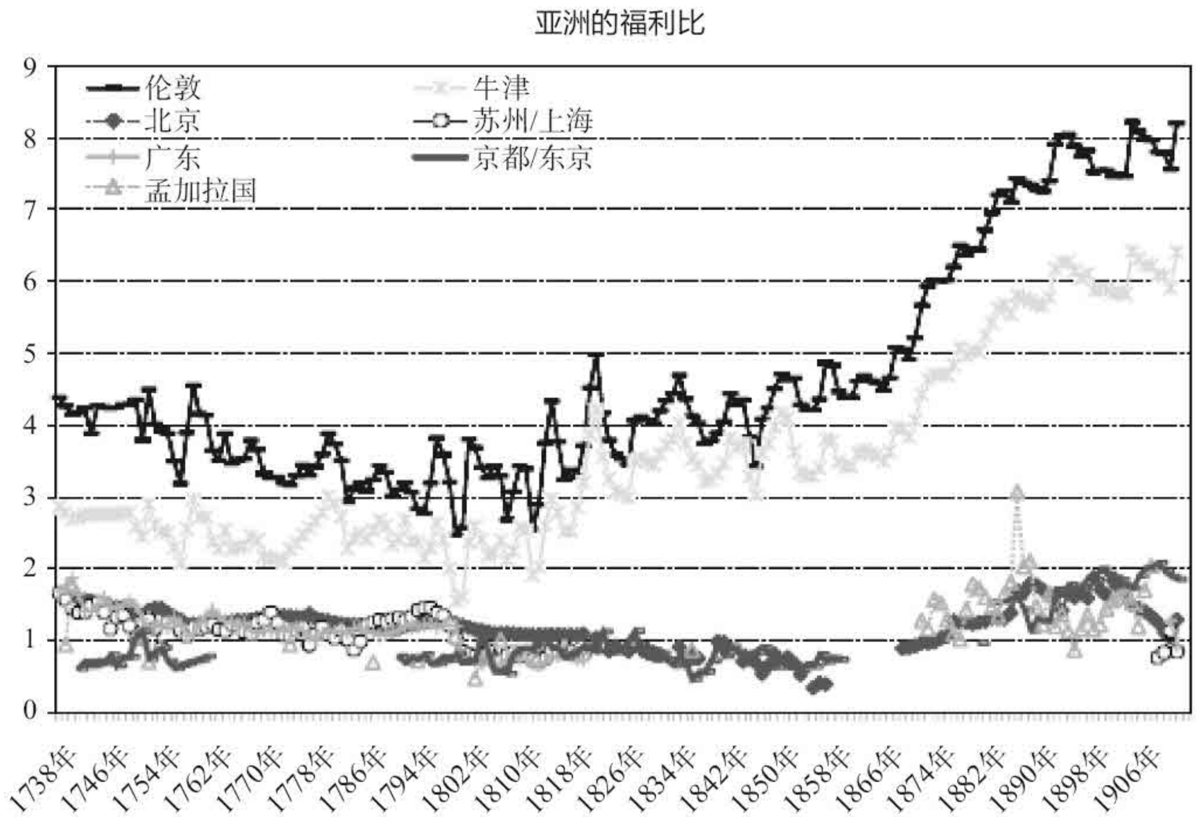
资料来源：艾伦，《农业生产力和农村收入》（*Agricultural Productivity and Rural Incomes*）

相形之下，乐观主义者们认为，1738~1900年的北京实际工资的确是稍稍上涨了（见图3.20），但仍然非常低，与落后的欧洲南部差不多，与兴旺发达的欧洲西北部相差甚远。1738年时，北京、上海、苏州和东京实际工资的购买力，还不及伦敦或阿姆斯特丹的一半，但大致与欧洲南部（如米兰）或中部（如莱比锡）相当。东方的工资实际上在1918年时仍与南欧非常接近，但在1820年的时候，中欧的工资就已经与东方拉开了距离，而向英国靠近了。

我们可以得出结论，1800年东方核心地带的能量获取水平低于西方核心地带，但低得并不多。按照我的计算，当时西方的能量获取水平是大约38 000千卡。在东方核心地带，农业产量很高，也有相当大量的煤用于取暖和做饭，但没有蒸汽动力，关于实际工资的资料也表明总体生活水平低于西北欧。所以，东方标准的能量获取水平在其核心地带（中国北方和沿海地区，再加上日本）大约为每人每天36 000

千卡，在不超过西方能量获取水平的前提下，不会过于高于这个水平；在不降至罗马帝国水平的前提下，也不会低得太多，与罗马帝国为伍似乎也是不可能的。

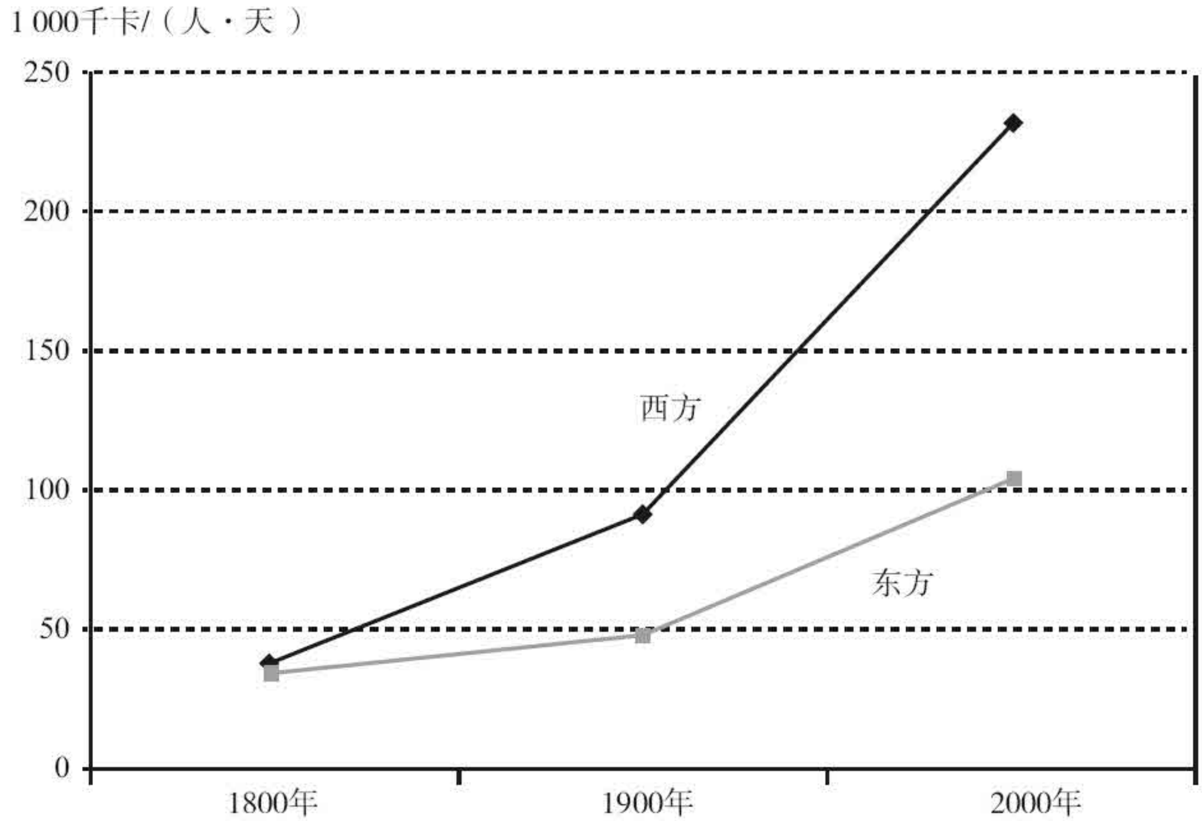
图3.20 公元1738~1918年欧洲和亚洲实际工资



资料来源：艾伦，《工资、价格和生活水平》（*Wages, Prices, and Living Standards*）

这些数字表明，东方核心地带的能量获取水平进入现代时期的时间（这么看来，大约是在1800年）仅比西方稍晚一点（见图3.21）。与传统/悲观论点相左的是，19世纪东方能量获取水平的确上升了，但增幅比西方要小得多。在19世纪全球力量重新分配中，西方占了先机，是由西方的腾飞，而不是东方的衰落决定的。同样，20世纪东方的全球地位提高，也不是因为西方的衰落，而是因为东方人学会了利用西方人率先利用的化石能量资源。

图3.21 1800~2000年东方和西方的能量获取情况



中国宋朝（960~1279年）

宋朝也许是中国前现代时期能量获取的巅峰时代。人口增长非常迅猛，从10世纪早期的大约5 000万人增至1200年的1.2亿人，但所有的迹象都表明，生活水平和能量获取水平上升得更快。

最清楚的文献证据来自需要大量燃料的冶金业。50年前，经济史学家郝若贝（Robert Hartwell）重新分析了宋朝税单，提出11世纪宋朝的铁产量比历史学家们先前认为的要高出20~40倍。据他计算，1078年时中国的总征税产量为75 000~150 000吨，比850年时增长了12倍。而且，郝若贝指出，1078年中国的铁产量大致比1640年的英格兰和威尔士高出2.5倍，超出1700年整个欧洲的产量的一半，大约与中国1930~1934年每年产量持平。郝若贝对文献的分析不断受到挑战，葛平德（Peter Golas）在《中国科学技术史》（*Science and Civilisation*）他

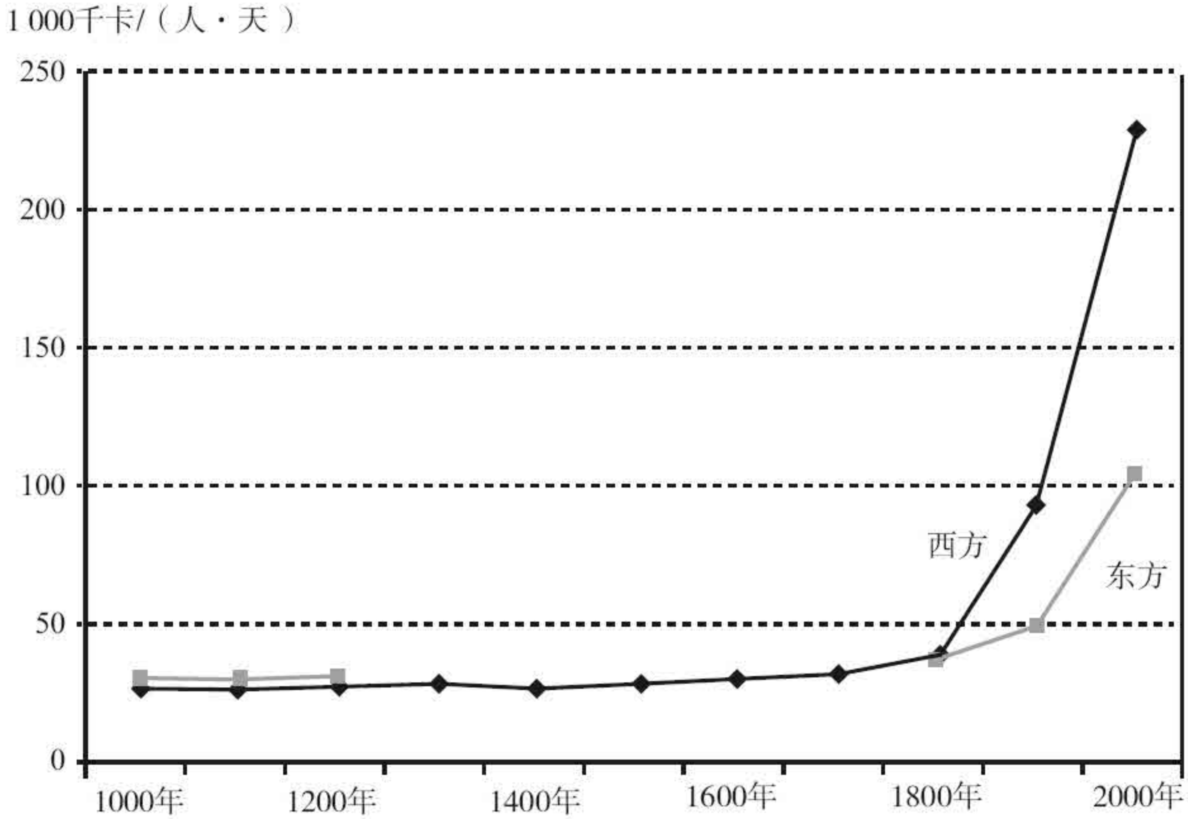
所负责的那卷中，提出郝若贝的铁产量数字存在整个数量级的偏差。然而后来，唐纳德·瓦格纳（Donald Wagner）在《中国科学技术史》他所撰写的那卷中得出结论，尽管郝若贝对不同文献的解读有缺陷，但他的数字肯定是大致正确的。中国历史学家漆侠独立研究后得出的结论是，宋朝铁制工具在农业生产中的大规模扩张，意味着11世纪时农民家庭对金属的需求一定达到了每年7万吨左右；而国家对铁制钱币和武器的需求还更大。铜的生产同样非同寻常，从997~1070年产量涨了4倍，从2 420吨增长到12 982吨，比1800年时全世界的产量都要高。11世纪和12世纪，中国金属制造业的副产品首次在格陵兰岛和南极洲的冰冠上留下了痕迹，就像1 000年前的罗马的银加工业。

郝若贝不断地将宋朝冶金业的扩张与1540~1640年的英格兰相比较，认为——像英国的例证一样——结果之一是炼铁极大地加快了以化石燃料代替木炭的进程。假如中国的铁器制造商们完全以木炭为燃料，那么1080年时他们就需要烧掉22 000棵成材的大树，远超过开封一带所能够供给的数量。所以取而代之的是，他们学会了用焦炭来炼铁，转而大规模地采煤。到1050年时，开采的煤产量之高，以致家庭做饭和取暖时烧煤比烧木柴都要便宜30%~50%。到1075年时，开封已经有了只出售煤的专门市场，1096年的官方文件也讨论了煤供应问题，甚至根本没有提及木材作为热源的事。对于在蒙古境内曾是大宋帝国边境的地区发现的铁制品和钢制品新近所做的分析，也证实了这一转变，在10~12世纪，冶金业用煤炭取代了木炭。

不幸的是，至今还没有发掘出的船只残骸、动物骨骼等的统计资料，无法与公元前900~公元800年的西方核心地带相比（见图3.5），但是来自文学、艺术及其他文物的质量证据证实了贸易、商业和制造业有巨大的扩张，纺纱机和水车得到了广泛应用。自20世纪80年代以来在广东沿海遭到大肆洗劫的大量宋朝沉船残骸，显示了船体在变大，船上货物越来越丰富。2007年得到完好打捞的“南海一号”沉船证实了这一点。

那时的房屋也变得更结实了，在12世纪的杭州，两层小楼非常普遍，与古老的中国城市形成了鲜明的反差。然而，大多数人也许仍住在只有一两间屋子的木质小房子里。有一些证据表明出售陶器和其他家居用品的大众市场在增长，但对家庭集合艺术品似乎没有任何统计研究。

图3.22 宋朝时期（1000~1200年）和现代时期（1800~2000年）东方能量获取情况，及与西方能量获取情况的对比



11~12世纪的能量获取水平（按照前现代时期的标准）无疑很高，但很难赋予其绝对值。铁制品的生产规模和中国的冰蕊污染记录显示，其能量获取水平大概和罗马帝国1 000年前曾达到的水平差不多（每人每天31 000千卡），或者达到了西欧大约公元1700年的水平（每人每天32 000千卡）。然而，中国没有发生任何我们可以称之为工业革命的事件，说明其能量获取不可能接近西欧1800年的水平（每人每天38 000千卡）。我姑且假定宋代的能量获取水平仍稍稍低于罗

马，在1100年时达到了每人每天30 000千卡，也许到1200年时稍稍超过了这个数字（见图3.22）。一个稍稍高于罗马水平，或许甚至与欧洲1700年的水平（32 000千卡）相当的数字，似乎同样是说得通的，但更高或者更低的数字——比如说达到了每人每天35 000千卡，或者低至每人每天25 000千卡——似乎就不大可能了。

现代早期的中国（1300~1700年）

20世纪六七十年代，经济史学家们普遍认为，中国在中世纪时期经历了生产力和生活水平的巨大飞跃后，在1400~1800年，农业和工业生产均陷入了停滞状态，然后在19世纪，由于内战、政府管理不善和西方帝国主义国家入侵的影响，实际上倒退了。

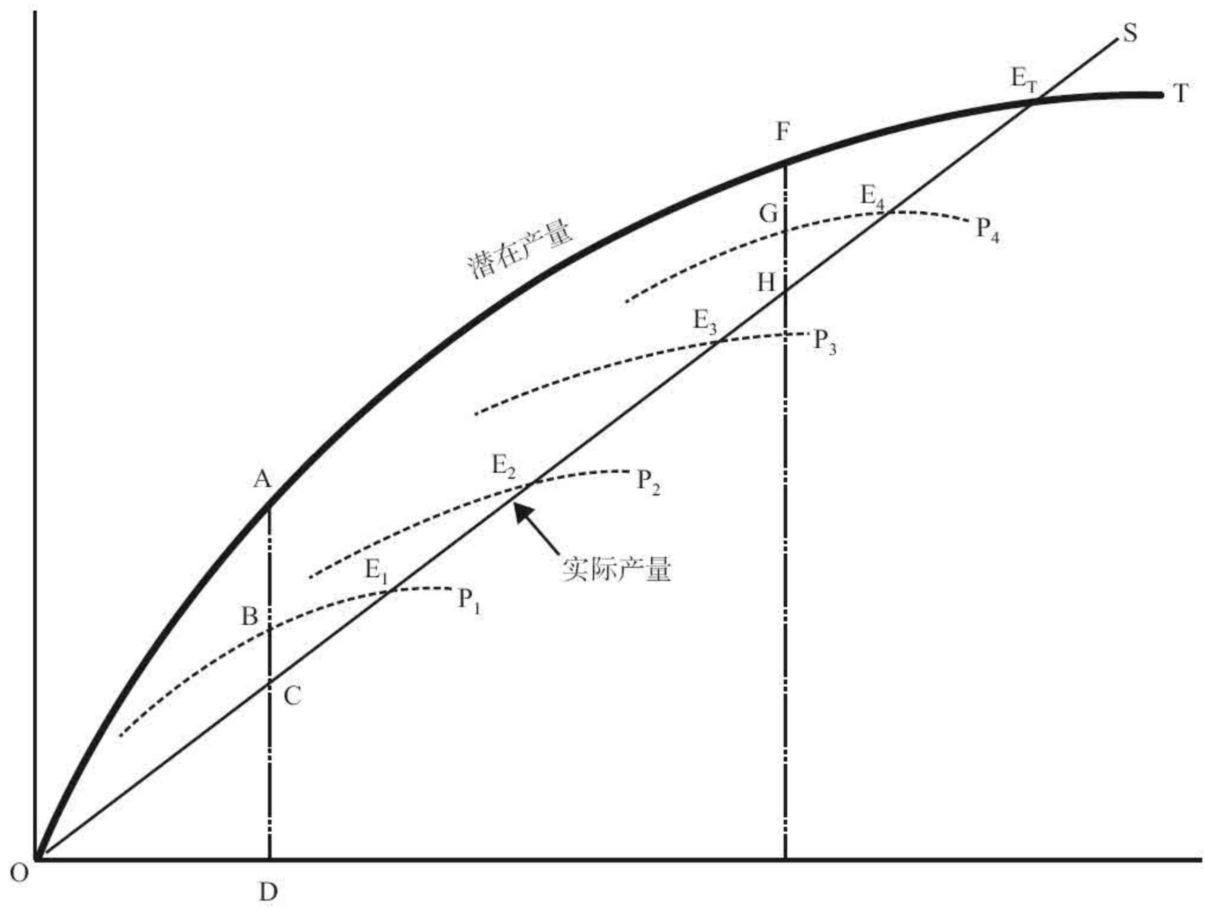
这种论点有好几个版本。德怀特·珀金斯在他对1368~1968年的农业产量的先驱性研究中，在卜凯（John Buck）对战争时期的研究的基础上，提出15~19世纪，最先进的农业技术从长江流域推广到中国北方，然后随清朝时期的移民，又推广到陕西，甚至西部更远的地方。据珀金斯的计算，长江三角洲地区水稻的产量到1300年时达到了极高的水平：为每公顷[注]3.5吨，是1800年英国地区产量（每公顷1.7吨）的两倍还要多，尽管按照劳动者人均产量计算，只有1800年英国水平的1/3（每公顷0.3吨比每公顷0.92吨）。中国的生产力与罗马帝国治下的埃及的灌溉小麦种植相比，成绩也极可观，埃及的产量可能大约为每公顷1.67吨和每人0.6吨。珀金斯认为，最先进的农业技术于1400年后在全中国普及，极大地提高了农业总产量，甚至人均产量也因为出现了较好的耕作方式而有所提高，但19世纪最优秀的农民仍不及14世纪最优秀的农民那么高产。

伊懋可提出了一个更宽泛的观点，中国的能量获取水平经过唐宋时期非同凡响的提升后，于14世纪陷入了一个“高水平均衡陷阱”（见图3.23），农业、工业、财政和交通运输业都达到了传统手段所能达

到的最高水平。伊懋可说，唯一提高生产力的办法就是转向化石燃料经济，但由于传统技术已经达到了如此完美的巅峰，东方没有诱因让人们进行那种能引发西方工业革命的革新。简而言之，这样的革新实际上会降低产量，因此受到人们排斥。

这两种论点都认为中国经济停滞了大约400年，这与20世纪中叶认为中国是永恒静态的传统西方理论是相符的。在同样的理念下，安格斯·麦迪森提出，1500~1820年，中国的人均GDP稳定地保持在600美元，只有1700年英国水平的一半。正如上面提到的，罗伯特·艾伦认为1738~1900年中国城市非熟练工人的工资相对稳定无变化，而长江三角洲的农业产量在1600~1800年略微下降了。

图3.23 高水平均衡陷阱

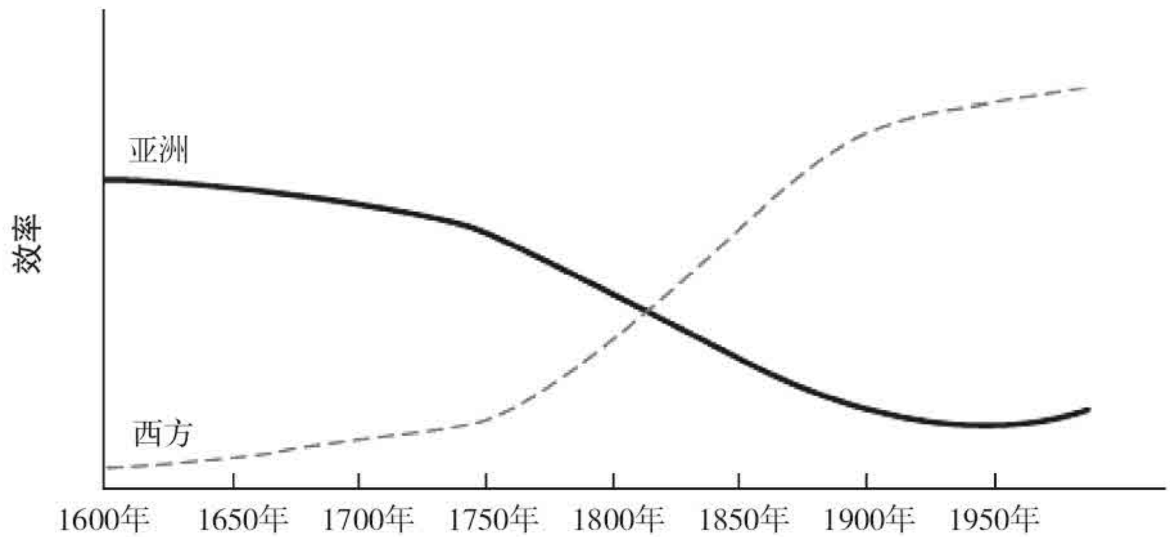


资料来源：伊懋可，《中国过去的模式》（*Pattern of the Chinese Past*）

自20世纪90年代，来自彭慕兰等人的挑战重启了争论。我本人的计算结果是，1200年宋朝的能量获取水平与罗马帝国非常近似（刚刚超过每人每天30 000千卡），而在1800年，能量获取水平只比当时的西方略低一点（每人每天36 000千卡）。这将意味着人均能量获取水平在1200~1700年增长了15%~20%。由于极少有历史学家对他们的早期现代时期生活水平提高的说法进行过量化分析，所以很难说我的数字是更接近于珀金斯/伊懋可/麦迪森/艾伦的观点，还是彭慕兰/王的观点。

然而，如果说1200~1700年的增长是一帆风顺的，似乎也不大可能。对长江三角洲一带新近的研究显示，有些地区这500年来的确很稳定，但总体而言，13世纪、14世纪和17世纪（也许稍好一点）都是饱受创伤的。尤其13世纪和14世纪出现了大规模的人口下降、城市毁灭和贸易崩溃。我不了解对经济具体领域的详细研究情况，但非常粗略地推测，能量获取水平在1200年左右达到巅峰——勉强超过每人每天30 000千卡——之后，可能下降了5%。譬如，1300年为每人每天30 000千卡，1400年为每人每天29 000千卡。这将导致1400~1800年是比传统模式更快的恢复时期。在3个世纪中，人均能量获取水平将提高20%。

图3.24 罗兹·墨菲关于1600~2000年西方崛起、东方衰退

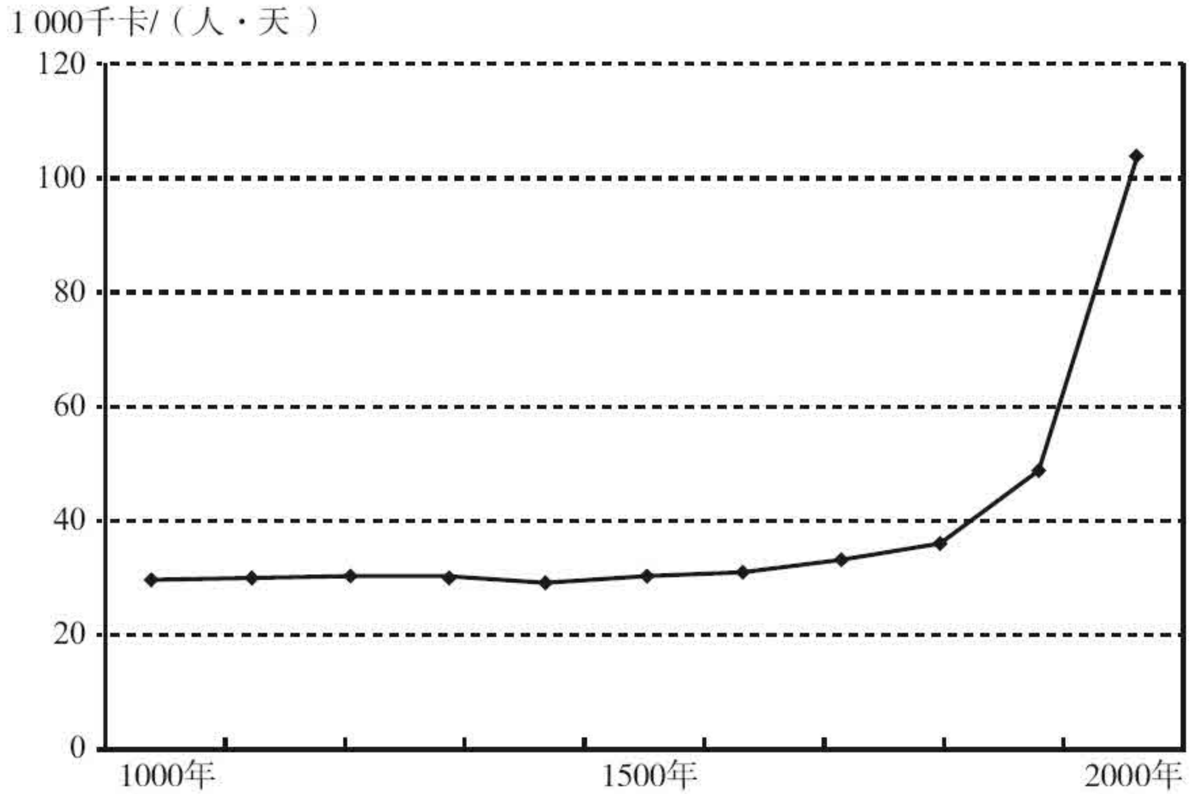


资料来源：罗兹·墨菲，《外来者》（*The Outsiders*）

未来的研究也许能够得出这些瞎猜的答案，但总体画面似乎是可信的：1200~1800年东方能量获取水平在稳定增长——实际上，按照前现代时期的标准，应当算是迅速增长；但西方能量获取水平的增长快得多。如果这样的说法是正确的，那么诸如安德烈·冈德·弗兰克（Andre Gunder Frank）和罗兹·墨菲（Rhoads Murphey）等历史学家们的主张——现代早期“东方的衰退”和现代早期“西方的崛起”对于促成19世纪西方的统治至少是同样重要的——必然是错误的，除非我们能发现证据，能证明公元1400年前东方的能量获取水平已经上升到足以与19世纪的西方相匹敌，然后又下降了——实际上，这似乎正是墨菲的图（y轴上没有数字）想展示的。

图3.25展示了我对公元第二个千年东方能量获取状况的估计。

图3.25 公元第二个千年的东方能量获取情况



汉帝国与罗马帝国（公元前200~公元200年）

古代中国在汉朝时期[公元前206~公元220年，通常被划分为西汉/前汉时期（公元前206~公元9年）、王莽改制时期（公元9~23年，也称为“新朝”）和东汉/后汉时期（公元25~220年）注]是个庞大而复杂的农业帝国，在很多方面都堪与同时代的罗马帝国相媲美。然而直到最近才开始有人将罗马帝国与汉帝国进行系统性的比较，而我们当前迫切需要考古证据的全面比较，如能以量化的形式更好。除非有这样的比较结果可供利用，否则这部分的估计必然是非常印象化的。

最普通的对汉朝经济的研究结果，几乎没有提供什么统计数字，但文献资料和对汉朝考古证据的定性分析，的确使尝试性的计算变得可行。汉朝最发达的农业在中国北方，特别是中原地区，但似乎明显不及当时最高产的罗马帝国先进。文献资料和考古发现都显示，公元

前1世纪时，中国即使最成熟的铁器加工业都无法与罗马帝国相比，而且在公元前第一个千年中，铁器只是非常缓慢地在中国农业中得到推广。公元前200年时，青铜工具、木制工具，甚至骨制工具和贝壳制工具，都比铁制工具要常见。关于犁的考古证据是有争议的，金属头的犁似乎到东汉时才得到普及。大规模使用犁牛和用砖衬井进行灌溉，似乎也是东汉而不是西汉的特色。文学作品也描述了一系列汉代出现的耕作方式改进，起始于公元前100年左右赵国的轮作法，但这些改进方法到底在多大范围内实施了，就不得而知了。很多最高产的技术和机械也许只局限于东汉的精英阶层使用。

给人形成的印象是汉朝的农业生产力不及罗马帝国，尤其不及尼罗河谷的灌溉农业发达。这也许不仅是印象。生产力在公元前200~公元100年之间肯定是增长了。贾思勰写于6世纪30年代的《齐民要术》一书，也表明农业技术（特别是水稻种植技术）即使在农业组织和基础设施遭到破坏的情况下也不断改善。许倬云搜集的文献资料表明，汉代农业是高度成熟的，但仍不及贾思勰时代的中国农业发达，也许也不及罗马帝国的农业发达。对汉代和罗马帝国成人身材的骨骼证据和营养的稳定同位素分析进行系统性的比较，将是极有助益的。

似乎没有任何全面的考古出土物目录，可供我们对罗马帝国和汉帝国定居点遗址的物质产品的丰富程度直接进行比较。对公元11年因黄河泛滥被淹没的一个村庄三杨庄新近的发掘记录的全面出版，将尤其富有价值。该村发现后，立刻被冠以“亚洲的庞贝”的称号。遗迹保存非常完好，考古学家们甚至发现了村民们在泥泞的田地中逃跑时留下的脚印。不大引人注目但却更具价值的是，至今可查阅的简要报告中描述的有着黏土屋顶的砖砌房屋，简直就像是稍小版本的同时代的罗马房屋。村民们的农具配备得很好，很多都是铁制的。从留存下来的黏土模型和布局证据判断，汉代城市里的房屋无疑是非常精美的，但总体的考古记录显示，中国的城市比罗马面积要小，结构也比较潦草。

文学作品中有大规模铁生产的描述。新近在朝鲜的发掘也发现了建于公元2世纪的令人惊叹的炼铁设施。沙伊德尔提出，罗马的货币供应大致是汉帝国的两倍，罗马最大财富值也是汉朝最大财富值的两倍。这些统计数字也许与人均能量获取水平只有松散的关联，但验证了人们认为西方古代时期能量获取水平高于东方的印象。汉代的能量获取水平似乎也低于宋代。至少从已公开的汉代考古证据来看，没有任何成就堪与宋代的煤铁使用、道路建设、技术革新、金融工具或者远程贸易水平相比。不过，与草原游牧部落和东南亚间的贸易在汉代的确有迅猛增长，而且如我在《西方将主宰多久》一书中所提到的，到公元2世纪时，汉帝国与罗马帝国之间可能存在直接的贸易接触。

依照目前的证据状况，任何关于汉朝能量获取的实际数字，都必然是猜测性的。汉朝的数字一定低于西方的巅峰罗马时代（每人每天31 000千卡）和东方的巅峰宋朝（每人每天约30 500千卡）。考古证据和文献记录也显示，汉朝在后罗马衰落的低谷期时（公元8世纪时为每人每天25 000千卡），能量获取水平高于西方，比西方晚期青铜时代的巅峰（公元前1300年时为每人每天21 500千卡）更是要高出许多。因此，汉朝的巅峰为公元1世纪，达到27 000千卡，到公元200年时由于组织和基础设施遭到破坏而稍有下跌（跌至每人每天26 000千卡）。西汉时期的增长似乎相当显著。这一时期能量获取水平提高了10%以上，从公元前200年的每人每天24 000千卡增长到公元前100年的25 500千卡；在公元前/公元1年和公元100年时，达到了巅峰水平：每人每天27 000千卡。正如上面所提到的，这些数字仍然是猜测性的，当有更好的比较考古资料可利用时，应当予以校正；然而，汉朝的巅峰水平似乎不可能低于每人每天25 000千卡，或高于每人每天29 000千卡。

图3.26 公元前200~公元200年和1000~2000年，古代、中古和现代时期东方能量获取状况，及与西方能量获取状况的对比

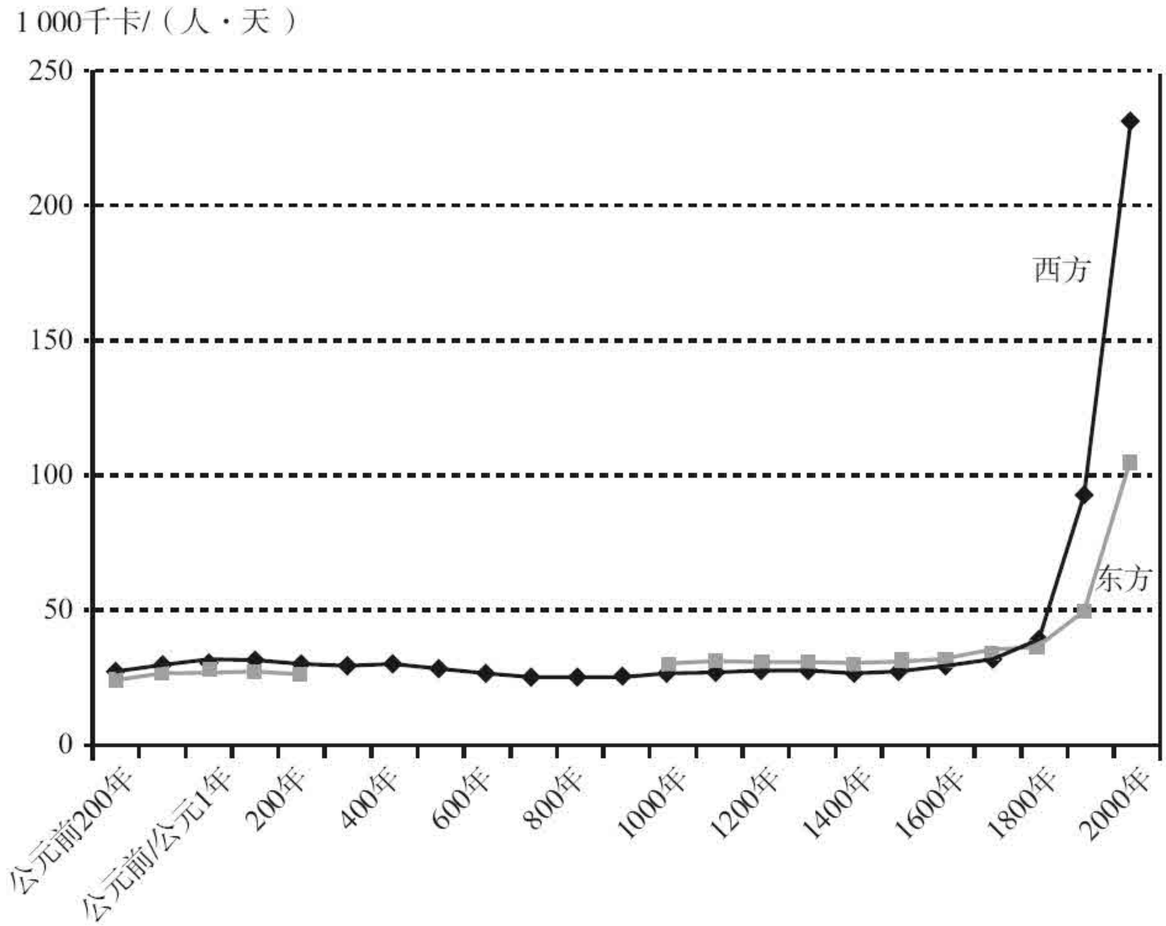


图3.26显示了对公元前200~公元200年和1000~2000年东方能量获取水平的估计，将其与过去2 200年西方分数的弧线对比，显示出西方核心地带在古代稍稍领先；在中古和现代早期，西方工业革命爆发之前，东方核心地带稍稍领先。

证据不足的800年（公元200~1000年）

“分裂时期”（公元220~589年）的能量获取历史甚至比汉朝的还要模糊。鲁威仪（Mark Lewis）新近发布了一份对这一时期的极富价值的研究成果，阿尔·迪安（Al Dien）也搜集了大量的考古资料，并做了同样有益的概括，但量化的研究仍然极其稀少。

像西方一样，中国经济的主要基础结构在公元200年后也瓦解了，尽管农业技术也许还有所改善。贾思勰的《齐民要术》展示了比任何汉朝文献都要详细的干谷种植知识，也披露了中国南方所实践的高深的水稻种植知识。看来自公元3世纪起，最好的水稻种植方法在长江以南得到稳步推广，到公元第一个千年末期时，水稻产量有了大幅提高。

随着5世纪轮船出现在长江上，原本只在佛寺使用的水车走进了许多家庭，地方特产如茶叶等得到广泛交易，经济的基础结构也在改善。国家强烈干预土地所有权，最著名的事例是“均田制”，但在4~6世纪那动荡的时代，那似乎有助于将农民拘系在土地上。在589年中国重新统一以及7世纪开凿大运河之前，汉朝以后的经济恢复很大程度上只局限于南方新种植水稻的地区，北方的商业衰落到连货币制度都在很大程度上消失了的地步；但是到公元650年时，整个帝国范围内的经济已经在复兴中了。水力灌溉得到了更广泛的应用；据记载，长安和其他一些大城市都出现了规模巨大的公共市场。755年安史之乱爆发后，国家政权的崩溃削弱了唐朝对经济的控制，但任何损失似乎都不及官僚干预的放松所获取的商业获益更大。

大多数历史学家似乎都同意，中国在公元600~1000年经历了经济的快速增长（依照前现代社会的标准），那段时期在经济上比西方先进。唐代上层社会的房屋至少像汉代一样显赫，佛教和宫廷艺术也很兴盛。然而，中国中古时期的考古学家们的注意力大都集中在艺术史和建筑上了，我们缺乏证据来定量分析这些变化对个人层面上的能量获取意味着什么。如果我估计的能量获取数字——公元200年时为26 000千卡，公元1000年时为每人每天将近30 000千卡——大致是准确的，那么中间的7个世纪增长率就大约为15%。上面所引用的资料给人的印象是，这些增长主要来自公元700~900年。我因此估计，在公元200~500年之间，能量获取水平大致保持在每人每天26 000千卡，然后在公元600年时升至每人每天27 000千卡，到公元800年时又升至每人

每天28 000千卡，公元900年时升至每人每天29 000千卡。图3.27显示了这些估计的分数，以及假如能量获取水平在公元200~1000年在平稳增长的情况下的分数（无论是算术级增长还是几何级增长），差别极微小。

图3.27 估计公元200~1000年东方能量获取水平的3种方法

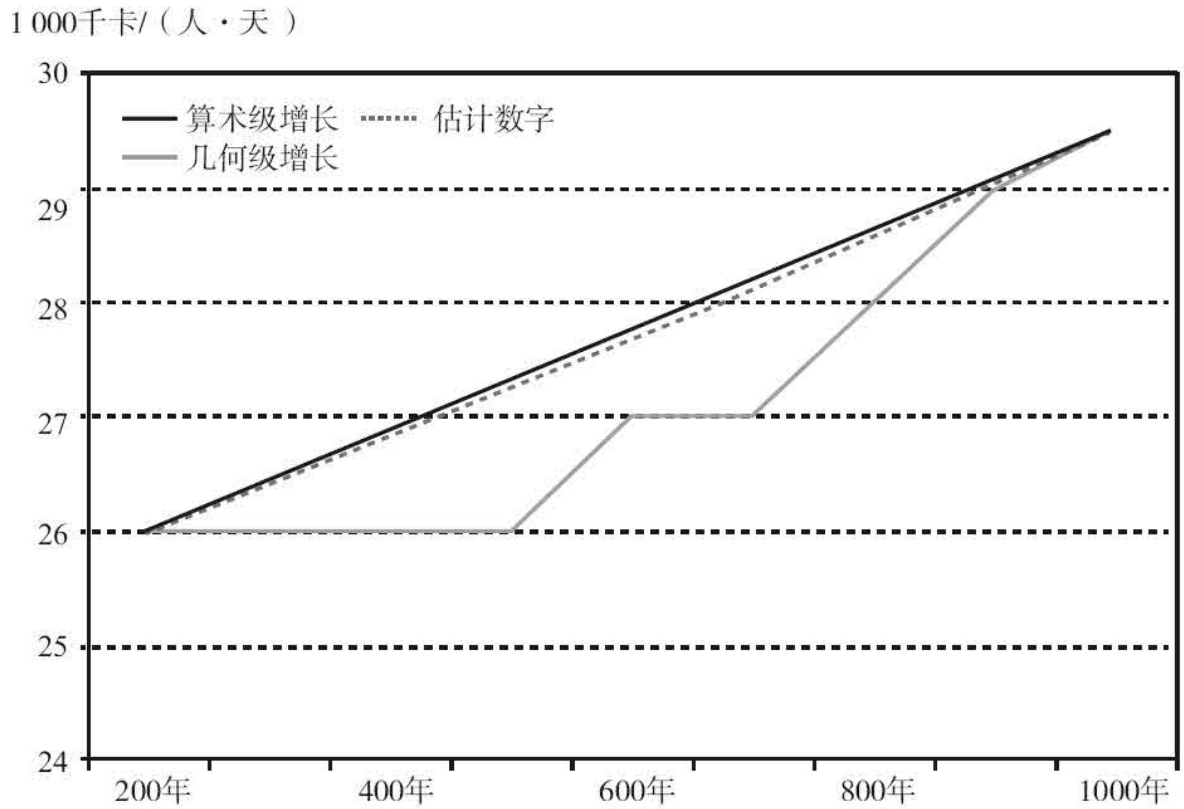
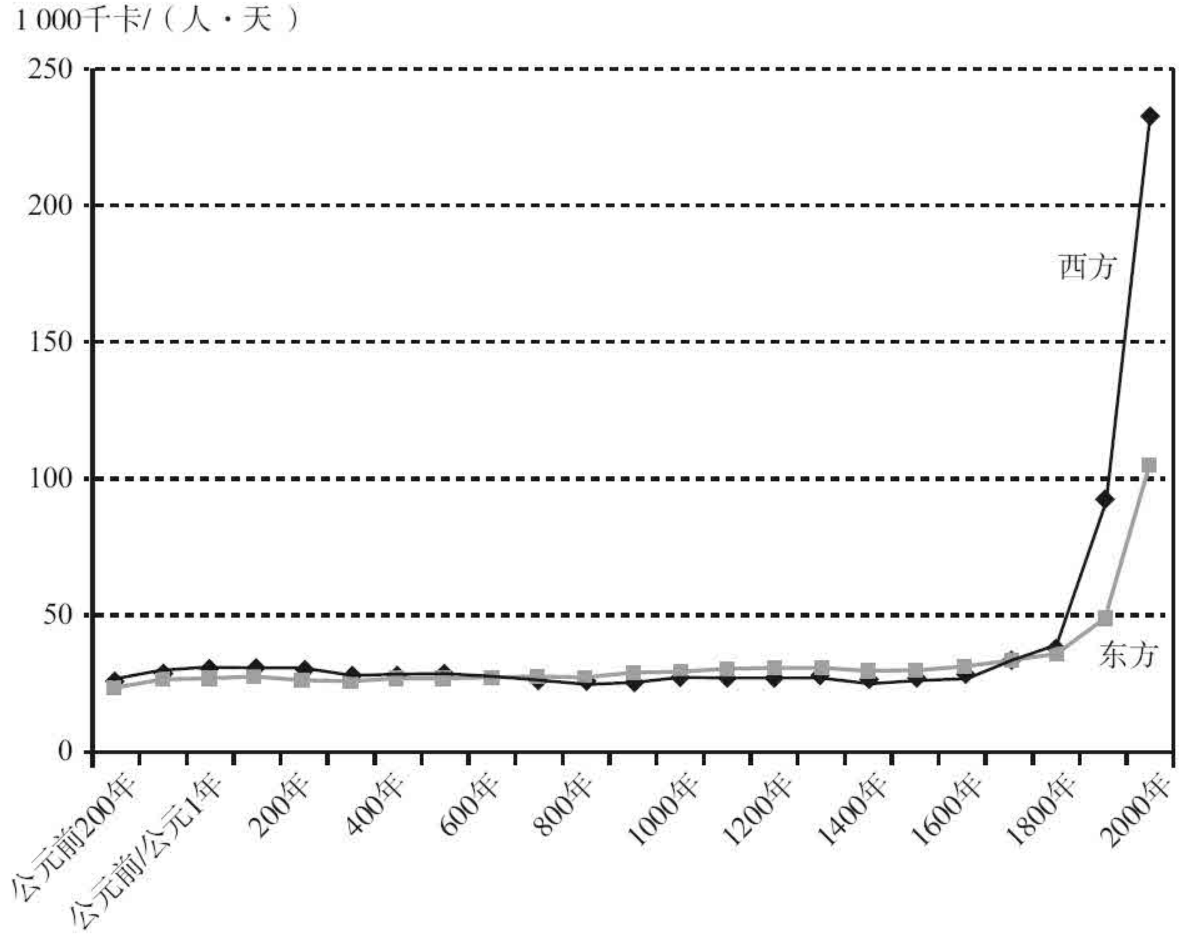


图3.28展示了我对自公元前200年以来整个时期东方和西方的估术。按照这些数字，东方的能量获取水平在公元563年，有史以来第一次超越了西方；不过，如果不是这样的话，公元1800年之前的2 000年的能量获取的历史就太平淡了。无论在欧亚大陆的东端还是西端，庞大的帝国都在冲击着有机经济可能达到的上限，但都无法突破。这是支持近些年关于欧亚文明的一个普遍看法——历史是循环的——的事实基础。在某种程度上，认为根本没有发生过什么重大变化的欧亚精英人士是正确的。

图3.28 公元前200~公元2000年东方和西方的能量获取情况



冰期晚期和后冰期时代的狩猎–采集社会（约公元前14000~前9500年）

我对冰期晚期和后冰期时代东方的能量获取水平的估计，很大程度上依赖于估计西方时所依赖的同样的灵长类动物能量学和人类进化研究成果。公元前14000年左右，东亚的智人一定已经能够获取大约4 000千卡的能量，否则他们就会灭绝；而假如他们获取的能量比这多得多，甚至能达到5 000千卡，那么我们就能从考古记录上看出这一点，比如会有更精致的建筑、物质文化，或者昂贵的食物热量。然而事实上，我们看到在将近5 000年的时间里，考古记录基本上没有什么变化。

在西方核心地带，能量获取水平在冰期结束之前就已经在增长了；但在东方，公元前9000年之前的考古遗址，完全没有任何建筑结构遗存。有一些证据表明，大约2.5万年前对动物尸骸的利用增多了。在中国南方的玉蟾岩遗址，也发现了手工制作、烧制温度很低的粗糙陶器——世界上最早的陶器，年代可回溯至大约公元前16000年。到公元前14000年时，中国北方和俄罗斯的远东地区也在制作陶器了。陶器的发明也许意味着人们开始食用需要煮沸的新型食物了，很可能是野生稻（在南方）和野生粟（在北方）。

然而，东方核心地带的情况与西方核心地带不同。在西方核心地带，到公元前11000年时，阿布胡赖拉的黑麦种子已经变得比较饱满了。而在东方核心地带，却几乎没有公元前14000~前9500年间人均获得的食物热量有所增长的较明确的证据。在吊桶环遗址，发现在公元前12000年时，野生稻谷被采集并带回了山洞，正好是在公元前10800~前9600年的新仙女木寒冷期之前，但野生稻谷在这个迷你冰期期间似乎又消失了，直到公元前9600年后才又重新出现。至今还没有发现在新仙女木期前有种植稻谷或其他植物的证据。这几千年当然会有其他变化，但似乎都是循环性的，而且规模极小无法度量。因此我估计公元前14000~前9400年整个时期，能量获取水平都在4 000千卡左右。

从采集者到帝国主义者（公元前9500~前200年）

正如图3.29清楚显示的，公元前14000~前9500年冰期后狩猎–采集社会能量获取水平合理有把握的估计数字（每人每天4 000千卡），与下一个估计数字（公元前200年西汉时期的每人每天24 000千卡）之间，还有很大的空白需要填补。我们可以为这9 300年简单地假设一个平稳的增长率，算术的或者是几何的，但实际上，将实地考古资料与文献数据相结合，借助经济人类学方法，再对比公元前200年之后的分数，我们是可以得到更准确的估计数字的（见图3.30）。

我将这一时期划分为3个阶段，先简要概括每个阶段的大致发展情况，再努力对这些变化对能量获取的意义进行定量分析。

图3.29 公元前14000~前9500年和公元前200~公元2000年东方能量获取状况

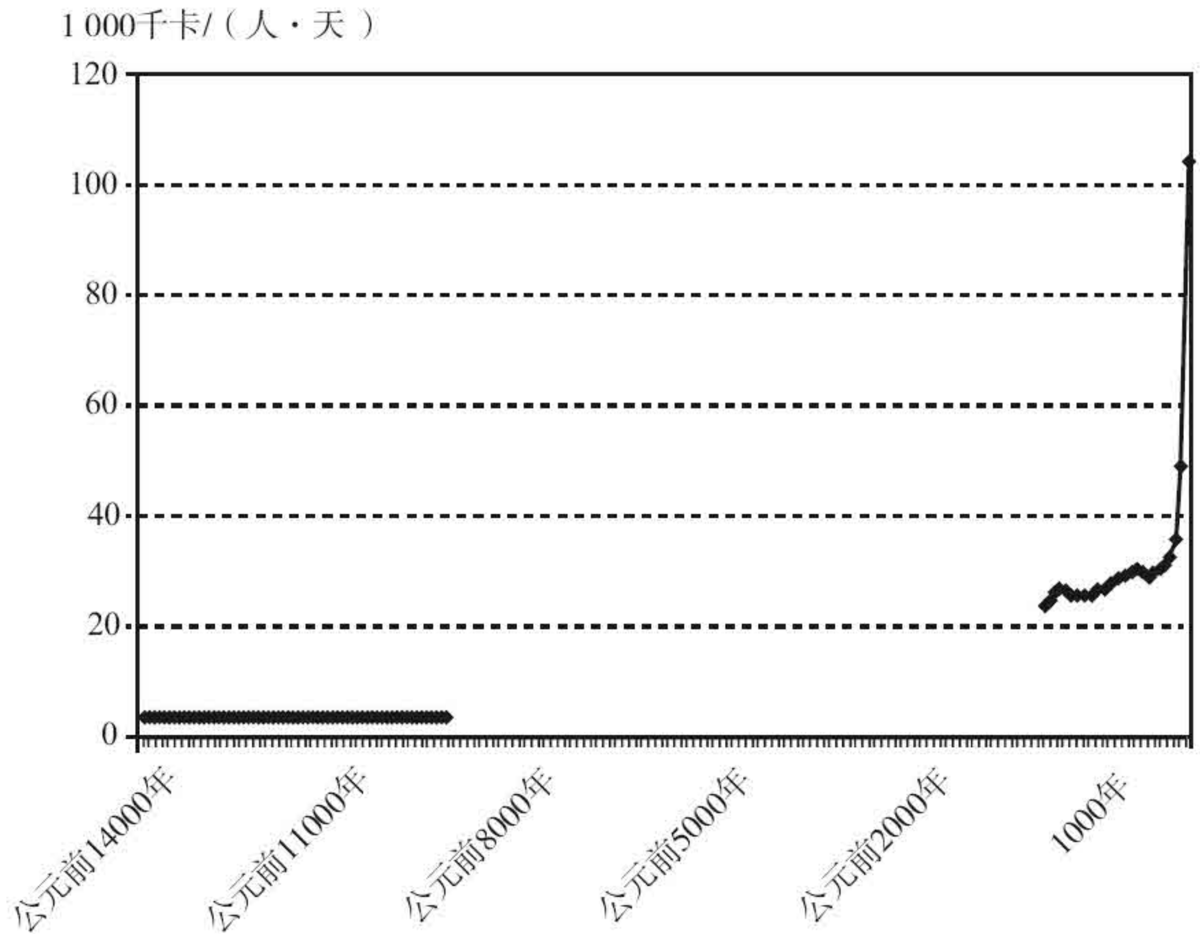
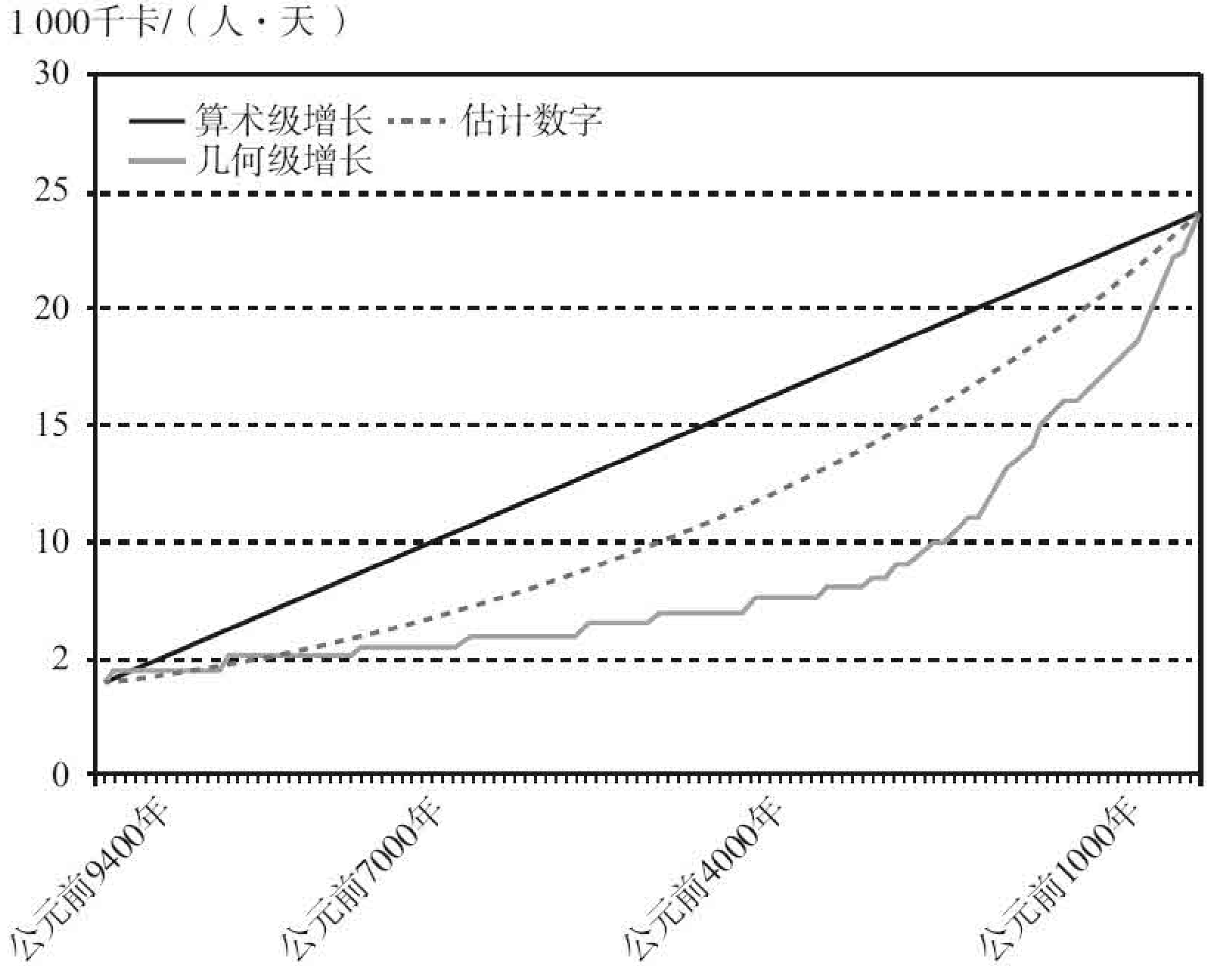


图3.30 估计公元前9500~前200年东方能量获取水平的3种方法



采集者和农民（公元前9500~前2500年）

在东亚工作的考古学家们经常热衷于将农业起源的时间尽可能地向前推。在中国北方的一些考古遗址，如早至公元前9000年或前8500年的河北南庄头遗址和虎头梁遗址发现的石磨和石辊，有时被视为小米种植的证据。在新近的一篇文章中，张家富等人提出，黄河河谷的龙王辿遗址出土的25 000年前的磨制石器，将中国农业的起源极大地前推至冰川时期。然而，对欧洲冰期磨制石器上的淀粉残渣的分析，证明其时间可回溯至公元前23000年，说明这些工具被用于将野生植物碾磨成糊状，制作前农业时代的粥或面包，中国的情况可能也是如此。新近在东胡林出土的公元前9000年的石磨上残存的淀粉，说明这些工具也是用于碾磨野生植物的，特别是橡子。迟至公元前6000年

时，在中国东北的白音长汗出土的石磨器上的淀粉中，橡子依然是主要成分。在渭河河谷，种植农业开始很久以后的很长一段时期内，野生食物也仍然是重要的食品。

东亚驯化植物的直接物证已成为激烈争论的话题。在大约公元前7000年的长江流域的彭头山遗址，制陶的胎泥中发现了稻壳灰屑，这肯定说明稻谷已经得到了驯化种植。自20世纪80年代以来，这已成为中国考古界的老生常谈。最近，蒋乐平和刘莉又提出，在长江三角洲的上山遗址和淮河流域的贾湖遗址发现的稻壳印痕和植物化石，证实了水稻被驯化种植是在公元前7000年左右。

然而，将中国的考古证据和论点与关于亚洲西南部农业的开端的争论相比较后，富勒（Fuller）等人提出，水稻一定是经历了很长时期的栽培，才形成了完全驯化的品种。他们认为蒋乐平和刘莉被不成熟的小穗所误导了，这些小穗在采集的野生稻中极其常见，上山遗址和贾湖遗址出土的都是野生稻。富勒等人的结论是，正经种植稻子是直到公元前5000年左右才开始的，也许是因为橡树覆盖面积减小，此前作为重要食物的橡子出现短缺。他们提出，完全被驯化的稻子是直到大约公元前4000年才进化而成的。他们还认为，中国北方驯化粟谷实际上先于南方驯化稻子，对于公元前5500年时栽种粟谷，和公元前4500年才出现驯化稻子，都有清晰的证据。

随后是激烈的交锋。情况经常是这样，似乎争论双方都有充分的根据：如果种植和驯化稻子像富勒坚称的那样晚，那么其在中国传播的一些特点就难以解释了；然而如果种植和驯化开始得像刘莉主张的那样早，此后一直缺乏较大的、明确无误的样品，也同样难以解释。进一步的工作无疑会解决这些问题。富勒提出的长期而持久的种植模式应当会得到证实，而传统观点可能也会被证明无误，即在水涝地上的公元前5000年前的河姆渡遗址发现的很多稻子都是驯化的，而到公

元前7000年时，贾湖、吊桶环、彭头山等地都已经出现了种植的稻子。

我们对东方核心地带中国的农业革命的想象，与西方核心地带亚洲西南部的现象越来越像了，只是大约晚开始了2 000年。正如在西方一样，似乎决定性的步骤并非发生在大河的河谷，而是在环绕着河谷的侧翼丘陵地区。农业革命的传播推广花费了数千年的时间，并且与模仿和移民相结合，还伴随着时间同样漫长的副产品革命。

这种趋势在中国可以通过农具的演变很好地看出来。例如，在6 000年前的半坡遗址，收割用的刀具占全部农具不到1/3，而到了5 000年前，在庙底沟遗址，收割刀具的比例就上升到1/2以上。在半坡遗址，锋利度不够的陶刀数量比石刀多，比例在2∶1以上；而在庙底沟遗址，石刀数量超过了陶刀。在半坡，（刀耕火种时伐倒树木所必需的）斧子数量比（在已经开垦的土地上翻土所必需的）铲子要多，比例在5∶1以上；而在庙底沟，锹的数量超过了斧子，比例在4∶1以上。庙底沟的锹的刃，也普遍比半坡的长50%（30厘米比20厘米），说明5 000年前的农民比6 000年前的农民翻土翻得更深，以改善土壤的通风条件。

也有其他证据支持副产品革命为时很漫长的设想，如中国北方新的稳定同位素分析显示，粟成为主要食物来源是直到公元前5000年以后的事情，长江流域也有证据表明驯化动物的过程也非常缓慢。

然而，东方和西方的巨大差异，是西方核心地带的耕种和驯化似乎开始得比东方核心地带早大约2 000年。即使我们忽略大约公元前11000~前10500年（显然早于新仙女木期）阿布胡赖拉种植的黑麦种子，到公元前9500年，就在新仙女木期刚刚结束时，种植的大麦和小麦已经确定无疑地出现在西方核心地带了。按照目前的证据状况，很难看到大约公元前7500年（甚至比这还晚，假设富勒是正确的）之前东方有种植稻子或粟的情况。完全驯化的小麦和大麦已经被确凿地认

定于公元前7500年左右出现在西方的侧翼丘陵区；驯化的粟直到公元前5500年前在东方都并不常见，稻子要到公元前4500年（或者按照富勒的说法，在公元前4000年）才被完全驯化。副产品革命在西方到公元前4000年时已基本完成，在东方到公元前3000年时还仍在进行中。例如，直到公元前2500年，我们才发现了真正令人信服的证据，可以说明东方典型的农业性别结构：男性从事户外活动，女性则从事室内活动。

像西方一样，东方食物热量总获取量的增长，是伴随着巨大的人口增长和缓慢而又惊人的人均非食物热量的增长的。已知最早的房屋时间可回溯至大约公元前8000年，在长江三角洲的上山遗址；更早的遗址只发现了灶台。房屋的面积在稳步增长，从公元前7000年的贾湖遗址平均4~6平方米的圆形的半地下的棚屋，到公元前4000年大河村遗址上30~40平方米的正方形的地面上的建筑。公元前7000年的贾湖遗址最大的房屋面积为10平方米，而公元前4000年的大地湾遗址却有150平方米和290平方米的“宫殿”，而且面积还只是根据房顶算出的。房屋的体积在公元前第三个千年以前增长也很缓慢，但在那之后，就突飞猛进了。

古代国家（约公元前2500~前800年）

公元前2500年以后，特别是公元前2000年后，随着更复杂的社会形态的出现，能量获取的增长率大大加速了。像在西方一样，由于没有对人体骨骼数据的大规模系统性的搜集和比较，未能直接记录下古代国家的出现对人类体格的影响，但存在其他变化的迹象。

其一是稻子种植在中国北方的推广，特别是在约公元前2300年之后；其二是动物骨骼在公元前第三个千年晚期和第二个千年的定居点中的大量增加。到公元前2000年时，驯化的猪一般在家养的畜类中占到2/3以上。文献记录也提到了公元前第一个千年农业组织方面的各种各样的改革，也许反映了真正的变化。

《孟子》第三篇上部提到井田制，据说是公元前第一个千年早期西周王朝设立的制度，尽管孟子所言一定是对当时混乱的现实的理想化描述。历史学家们经常把这种土地占有制度称为封建制，尽管这似乎不完全适当。

总的来说，东方农业生产力（无论是劳动力人均还是每单位土地）似乎一直比当时的西方要低得多。公元前第三个千年的考古遗址中出现了不少铜器（大多是装饰品），但在公元前800年之前，极少见到金属农具的实例。在公元800年前，木头、石头、骨头和贝壳始终是占绝对优势的最重要的农具材料。除非有更好的证据出现，否则我们不得不得出这样的结论：东方古代国家农业产量的增长，慢于美索不达米亚和埃及的西方古代国家的灌溉农业体系所带来的。

然而，在公元前2500~前800年，非食物能量获取水平看来的确有强劲的增长。由于考古学家们让人出乎意料地对后新石器时代的中国定居点缺乏兴趣（不是因为缺乏实际的遗迹，而是因为考古学家们只关注贵族坟墓和重大建筑），我们的想象受到了限制。为数不多的出土物的确显示，到公元前800年时，房屋的面积和质量都有所改善。洞穴屋依然在建造，但更多人住到了地面上，有人住在结实的矩形房屋中，有挖了防护壕的地基、夯土或泥砖的墙、灰泥抹的地面，以及墙围。一些房屋中有绘画装饰，也有一些房屋是围绕着一个宽敞的院子建造的。墓穴中发现的浸满水的木器，说明细木工技术得到了较大改善。这些事物发展的年代顺序尚不清楚，但宽泛地说，我们可以相信房屋建造水平在公元前第三个千年的晚期到第一个千年的早期，有了长足进步。

家居用品的数量和质量也有提高。公元前第二个千年时，陶器已经普遍使用飞快旋转的轮子制作了，丝绸、漆器和玉器也更加常见了。最早的铜器出现在大约公元前3000年，几乎可以肯定是受到了越过大草原传来的西方冶金技术的推动。金属似乎非常罕见，直到公元

前第二个千年早期，二里头、郑州和安阳才出现了巨大的铸造作坊，打造武器和一些手工工具，最重要的是铸造祭祀器皿。铜陵保存完好的矿井遗址显示了早至公元前1600年的中国冶金的规模。

鲜有像样的家庭考古资料，意味着我们对金属的日常使用情况知之甚少，不过坟墓中出土的物品和财宝似乎意味着，到公元前800年时，青铜器皿已在某种程度上照社会等级自上而下流传了。在上层社会，金属的使用量是极大的，公元前12世纪的后母戊鼎（大概是从安阳的王室陵墓中发现的），使用了将近一吨的青铜。自公元前1046年商/周更迭后，有铭刻的青铜器皿的数量呈爆炸式增长，也许证实了一个非常富有的贵族阶层的出现。考古学家们也已经确定了上层社会中发生了“祭祀革命”，自公元前9世纪起开始用青铜器陪葬，这似乎与青铜工艺的巨大进步，包括脱蜡法和焊接法的应用，是相适应的。

上层社会的重大建筑在公元前2500年以后也有重大发展。公元前第三个千年晚期最大的一些遗址（有时会占地200~300公顷）开始夯出通常会高达两米的土台。其中最宏伟的陶寺遗址有早至公元前2600年的一个占地5公顷的宫廷院落；到公元前2300年时又筑起了厚达9米的防护墙，有巨大的圆形碑，还有墙上有绘画的宫殿。

大约从公元前1900年开始，在可能是夏朝和商朝的都城的二里头和郑州，建起了规模大得多的宫殿。而从公元前13~前11世纪，尽管曾遭受洗劫，安阳的商朝王陵无论依什么标准，仍然会给人留下深刻印象。至今发掘出的西周宫殿都不及其前朝商朝的宏伟，不过其都城酆京遗存的房屋仍然极大而坚牢。公元前1046年后，富裕的墓葬骤然激增。贵族炫富的程度和能量获取水平在公元前1000~前800年也许是持平的，但比公元前2500年时要高得多。

像西方一样，在古代国家时期，确定无疑地发现了最早的地区性崩溃的证据，最明显的是公元前2300年左右陶寺的陷落和山东复杂社会的瓦解。不过，像公元前2200年和公元前1750年西方的崩溃一样，

陶寺/山东的衰落对于能量获取水平并没有产生明显的影响，至少对于如此粗略的度量工作来说是如此。

春秋/战国时期（公元前800~前200年）

东方没有经历像公元前1200年西方的崩溃那样的灾难。西方的崩溃使其核心地带的能量获取水平一连下降了好几个世纪。东方的能量获取水平则相反，上升得越来越快。像古代国家时期一样，缺乏对人体骨骼资料的系统性搜集和家庭考古的不足，限制了我们，但相同的是，仍有充分的证据使我们可以勾勒出总体的画面。

文献资料证实，土地占有制发生了进一步的变化，特别是出现了向私人土地所有制的转变。土地掌握在法律上自由的农民手中，国家课税取代了有人身依附关系的农民在土地上为他们的主人劳作的制度。这方面最早的清楚记载是公元前594年鲁国对农产品课税。到公元前3世纪时，这种向自由持有土地的制度的转变可能已经完成。这种土地所有权的改变大概鼓励了农民自发地加大了投入；如果真是这样的话，更高的产量便是其理所当然的结果。土地所有制的改革还伴随着农业理论和方法形成了成熟的文献并不断发展，以公元前440年魏国的李悝为开端。

多茬复种的文献证据似乎也是与这种新的土地所有制相伴的。到公元前200年时，两种作物轮种（中国北方是小麦和粟，南方是粟和水稻）已显然成为常态，偶尔也会种些豆类，可能两年轮种3种作物。还有一些历史学家提出，“牛”字的流传说明公元前第一个千年中叶时，役畜也变得重要起来（至少是在上层社会）。然而，在更确定的文献证据的基础上，我们认为国家大规模地涉入水利工程始于公元前4世纪30年代魏国的地方官西门豹。战国时期所有国家都大力投资开凿运河，以提高农业产量，其最高峰为大约公元前300年秦国的李冰在新征服的四川兴建的大型水利工程。

金属工具最早开始大规模使用大概是在公元前800年以后。李学勤和唐纳德·瓦格纳都认为，公元前800~前500年，在长江下游地区，青铜工具变得越来越重要，但一些考古学家仍对此持怀疑态度。然而，到公元前500年时，中国也开始使用铁器（可能像青铜技术一样，制铁工艺最初也是跨越草原由西方传来的）。中国的铁匠进步非常迅速，在公元前6世纪时就炼出了真正的钢，公元前5世纪时铸出了生铁（而欧洲的铁匠是直到公元14世纪才掌握了这项技术）。到公元前200年时，铁制武器开始取代青铜武器，铁制工具的使用显然也越来越普遍。不过，青铜业依然繁荣。公元前6世纪的铜绿山铜矿显示了异常成熟的建筑水平，有设有树木线的竖井。侯马的铸造作坊也同样能给人留下深刻印象。

这段时期，商业也迅猛发展。公元前625年从鲁国的臧文仲开始，大臣们在各自国内巡视以废除关卡。诸侯国都必须向商人们保证不干涉贸易，水运也变得越来越便利。中国商人在没有受到西方发展影响的情况下，于公元前5世纪开始独立铸造和使用青铜钱币。到公元前200年时，已有数以百万计的青铜钱币在流通。中国考古学家还没有像西方考古学家那样对船只残骸、动物骨骼、碑铭和铅污染状况进行量化，但在公元前800~前200年之间贸易有了巨大增长，似乎是非常清楚的事实。

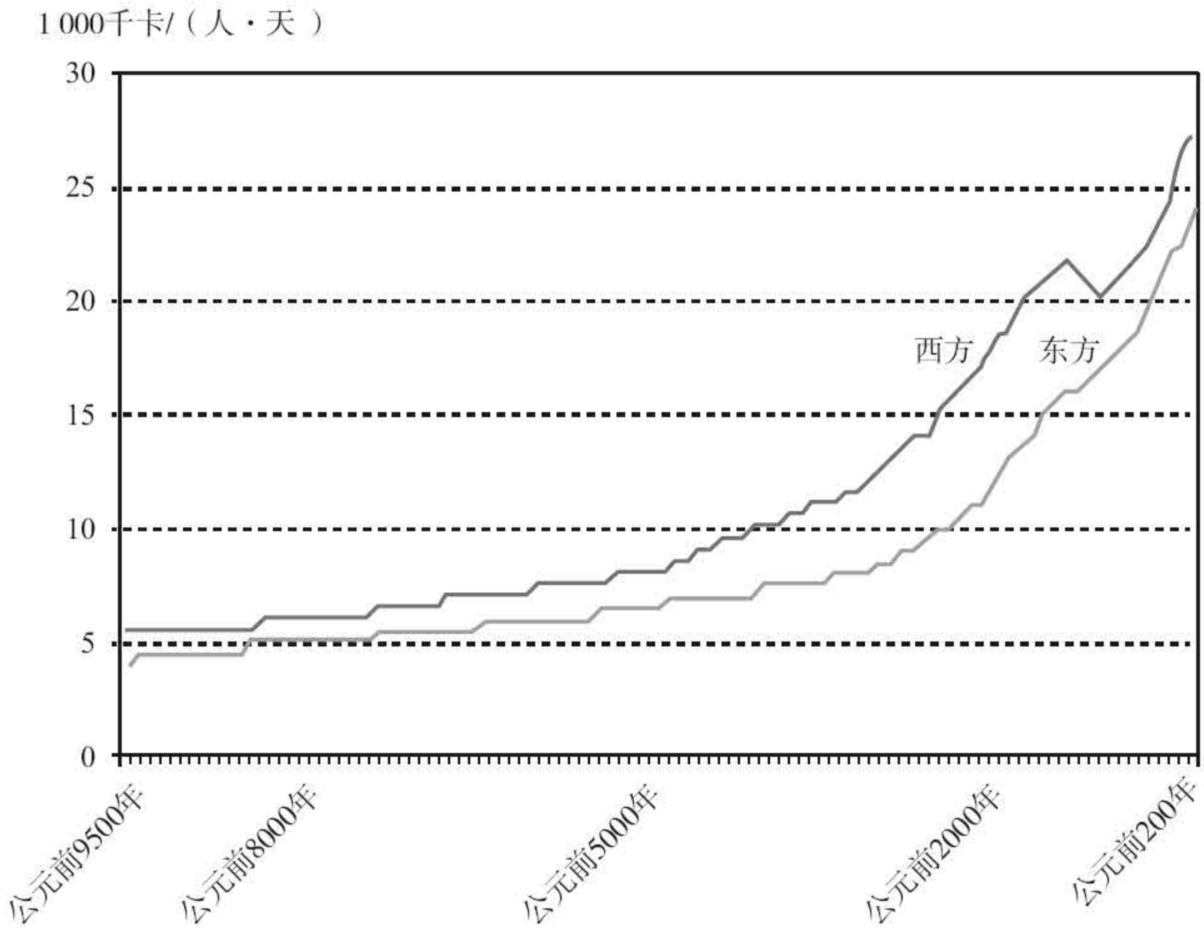
计算分数

图3.30显示了3种不同的办法，来填补公元前9500~前200年能量获取分数估计的鸿沟，或者索性假设能量获取以算术或几何的规模稳步增长，或者依据实际证据进行估计。算术级增长似乎不大可能：图3.30中上面那条线意味着，自大约公元前7000年贾湖文化建立到大约公元前5000年河姆渡文化建立，和自大约公元前2300年陶寺文化毁灭到大约公元前300年秦国在四川兴修水利工程，能量获取水平的增长是

一样快的。这不可能是正确的。我们或许应当假设能量获取水平的增长是几何级的，指数随时间而增长。

我们在将西方的考古数据转化为消费水平时所遇到的所有挑战，在东方进行转换时也都会遇到，但将东方的考古发现与西方相比较，显示出东方实际上是沿着一条与西方非常相似的轨道前进的。主要的不同在于，东方走上耕种和驯化的道路比西方晚了大约2 000年，其能量获取水平也相应落后于西方。在从采集向农业过渡的时代，最初差距似乎也停留在大约2 000年。从公元前9500~前6000年，东方的能量获取水平增长了大约50%，从每人每天4 000千卡增至6 000千卡；到公元前2500年时，随着副产品革命的完成，又增长了50%，增至每人每天9 000千卡。在这个时候，也就是埃及人建设大金字塔的时代，东方的能量获取水平似乎还停留在约公元前4500年时西方核心地带的水平，当时西方最早的大城市，比如布拉克丘和苏萨，都已经出现了（见图3.31）。

图3.31 公元前9500~前200年东方和西方能量获取状况



不过，公元前2500年后，东方能量获取水平的增长要快得多。由于东方的数据实在太少，我们只能就总体印象而言。在我看来，到公元前2000年埃利都腾飞前夕时，东方的能量获取水平一定大致达到了大约公元前3500年西方核心地带的水平，即苏萨时代和乌鲁克扩张的前夕（每人每天11 000千卡）。公元前1500年，当商朝建设郑州时，东方的能量获取水平与公元前2400年的西方的水平大致相当了，即乌尔王陵和埃及大金字塔兴建的时期（14 000千卡）。令我惊讶的是，到公元前1000年，当周朝取代了商朝时，东方能量获取水平已经与同时期的西方核心地带不相上下了。当时西方正处于危机后的恢复时期，埃及的中王国取代了旧王国，美索不达米亚的乌尔第三王朝取代了阿卡德王朝的城邦（每人每天17 000千卡）。不过，到公元前500年时，因为西方在公元前1200年左右崩溃，并且恢复缓慢，东西方差距

进一步缩小了。到公元前500年时，东方能量获取水平已经和东西方大约公元前800年的西方差不多了。当时亚述帝国正面临巨大危机，不得不进行上层体制改革——当时的水平是21 000千卡，当然也是西方于大约公元前1400年时达到的水平，即埃赫那吞和奈费尔提蒂开始奇特的宗教和政治实验之前半个世纪。

这些估计当然还需要更好的证据支持（无论是西方还是东方），目前都还只是推测而已。但如果这些大致在正确的范围内，那就意味着东方的能量获取水平在大约公元前9500~前3500年这6 000年的时间里大致翻了一番后，在公元前3300~前1300年这2 000年的时间里又翻了一番，然后在公元前1300~前200年这1 100年间又增长了50%。

公元前1200年左右西方崩溃是到公元前200年时，东西方差距缩小到只有300年的主要原因，但差距缩小的趋势早在此之前就开始了。在公元前2200~前1200年的1 000年间，实际上西方能量获取水平只增长了31%，而东方增长了52%。为什么会发生这种情况，原因还不十分清楚，尽管目前看来，东方的青铜技术很可能是向西方学习来的，并且驯化的小麦也同样是通过跨越大草原的商人获得的。这是否是东方迎头赶上的唯一原因，塔里木盆地里那些保存完好的干尸是否就是把很多技术从西方传到东方的中亚商人，是否还有至今未知的其他原因促使东方社会在古代国家时期比西方发展得快，都还有待研究。

无法突破的“硬天花板”

图2.5显示了过去16 000年间的能量获取状况，也展示了《西方将主宰多久》一书中的主要论点。社会发展指数的其他维度——社会组织（通过城市规模来度量）、战争能力和信息技术——终究不过是使用能量的方式；尽管单纯度量能量获取无法涵盖社会发展的完整内

涵，但能量必然是任何指数的中心支柱。因此我讨论能量获取的证据时比另外3个特点更为详细。

显然，还有许多工作需要做。尽管能量获取是历史的脊梁，我们关于能量获取的证据却是零散和不尽准确的。总体而言，西方历史比东方历史有更多的资料可供研究，西方历史的确有可供量化的证据存在。像史前考古的情况一样，致力于西方考古的学者们对于发掘结果，通常会比东方的同行们做出更多的综合整理。特别是西方学者们做了更多的家庭考古工作，并对实际工资进行了更多的研究。

随着证据基础的改善，新发现会解决这里遗留的一些问题。例如，也许有一天我们能够更有信心地断言，西方能量获取水平在罗马时代的巅峰期究竟是公元前1世纪还是公元1或2世纪，其水平是否当真比东方的巅峰期宋朝要高（以及东方的巅峰期是否真的是12世纪）。我们应当也能够证明，在公元第一个千年的早期到中期，东方和西方的能量获取水平是否都真的下降了；公元前1200年左右，西方的危机是否真的拉低了能量获取水平（我认为是的）；公元前2200年左右西方的危机和公元前2300年左右东方的危机，是否也拉低了能量获取水平（我认为没有）。更好的证据将不可避免地加强我得出的一些结论，而削弱我的另一些结论。

不过，总体情况显然是有充分坚实的基础的，即使所有具体分数都还存在争论的空间。能量获取水平在最后一个冰期末期时非常低，不会过分高于每人每天4 000千卡，增长得也极其缓慢。食物热量有所收益，但是，正如马尔萨斯在两个世纪前所认识到的，这些收益通常都被转化为更多的人口，增加的人口消费了增加的收益，使得大多数人的能量消费水平仍低于每人每天2 000千卡。但是，也如马尔萨斯所看到的，非食物热量也有大量收益，而且随时间积累。因此，总体（食物+非食物）能量获取水平是几何级增长而非算术级增长的，指数是随时间推移而增长的。

无论在东方还是在西方，我们都能比较清楚地看到耕作开始的大致时间（西方为约公元前9500年，东方为约公元前7500年）、驯化动植物开始的大致时间（西方为约公元前7500年，东方为约公元前5500年）、古代国家兴起的时间（西方为约公元前3500年，东方为约公元前2000年）、大帝国创建的时间（西方为约公元前750年，东方为约公元前300年），以及最重要的化石燃料工业兴起的时间（西方为约公元1800年，东方为公元1900年）。从古代大帝国的巅峰时期到工业革命兴起，大约2 000年的时间里，能量获取水平一直被控制在一个我称之为“硬天花板”的上限之下，略高于每人每天30 000千卡。这标志着农业社会所能达到的极限就是如此。这也在很大程度上解释了从古代和中古时期流传下来的精英作品中普遍存在的人性达到其顶峰、历史是循环性的、最好的时代在过去等感觉——正如自公元1700年以来西方能量获取水平爆炸式的增长，在很大程度上解释了18~19世纪众多欧洲思想家，以及20~21世纪的美国人抱持的乐观主义。

---

1. 1公顷=10 000平方米。——编者注
2. 应还包括更始帝时期（23~25年）。——编者注

The Measure of Civilization

How Social Development Decides the Fate of Nations

第四章

社会组织：上限未定

办法、假设和来源

社会科学的一个长期的研究传统，特别是在考古学、人类学、经济学和城市研究领域，在于表明一个社会中最大定居点的规模与其社会组织的复杂程度有着极其密切的关系。这种关联性远非完美，但对于探索跨度为16 000年的社会发展指数这种粗线条的研究，已经足够了。

原则上讲，探究城市规模还有一大好处，就是概念上比较简单。我们所需要做的只是：（1）确定东方和西方在历史上每个需要计算指数值的时间点上的最大定居点的规模；（2）确定公元2000年时世界上最大的城市规模；（3）将公元2000年时最大城市的人口数除以250（社会发展指数所能赋予“社会组织”这个特性的满分）；（4）再将过去城市的人口数除以那个数字。

由于对城市边界的定义和统计数据的可靠性的看法不同，人口统计学中对于公元2000年世界上最大城市的规模，意见很不一致。为了设立一个相对无争议的基线，我索性采用英国《经济学人》（*Economist*）杂志社出版的《袖珍数字世界》（*Pocket World in Figures*）中的估计：东京高居榜首，其人口为2 640万；纽约则是西方核心地带的最大城市，人口为1 670万。当然，还有大量其他估计数字可供我使用，但所有可靠的数字似乎都与上述数字相差不远。

这个起点意味着东方在公元2000年的社会组织方面得到满分250分，每106 800人得1分。纽约有1 670万居民，因而为2000年的西方挣得156.37分。值得记录的最小分值为0.01分，只需1 000人即可获得。这意味着，与能量获取分数不同的是，社会组织分数的确可能降至0分，公元前4000年之前的东方和公元前7500年之前的西方，人口都太少而无法度量。

以这种方法计算组织分数，主要的障碍是经验主义。对于早期定居点，我们必须借鉴考古学和人种志/历史的类比。估算极其依赖于对定居点面积的度量和根据有记录的人口密度的外推。人类学家罗兰·弗莱彻（Roland Fletcher）曾向我们展示了人口密度统计的差异有多大，尽管这些数字似乎都遵循了一般原则。在某些情况下，例如古希腊，估计数字可能非常可靠，误差幅度极小；而在另一些情况下，比如公元前第三个和第二个千年的美索不达米亚，估计数字就不那么可靠了。总体而言，大多数记录良好的前现代城市，很少有人口密度超过每公顷200人的，每公顷将近100人的城市更为常见。偶尔也可能有前现代城市的人口密度高达每公顷500人，但这样的人口密度极其异常，需要有非常清楚的证据。然而，20世纪和21世纪的超大城市中极小的村庄和特定地区，有时人口密度却完全可能超过每公顷500人。

从古代起，就有一些文学作品提到城市规模，但这些资料往往不可靠，因为古代城市的居民通常并不知道自己周围住了多少人。这意味着在现代之前，考古和类比仍然非常重要——尽管由于同时代没有很像前现代时期的“城市巨人”罗马和长安那样的城市，使得过去3 000年的类比比史前时期的还成问题。在近代，食物进口的资料有时也会留存下来，这给了我们又一个核实人口规模的办法。而在现代，我们可以利用相当准确的官方统计资料。

一些作家提供了有精确数字的关于城市历史的概述。特蒂斯·钱德勒（Tertius Chandler）所著《4 000年城市发展史》（*Four Thousand Years of Urban Growth*）是一部价值无量的参考书，可能也是被引用最多的著作，尽管书中提供的资料来源太少。其较早的版本，《3 000年城市发展史》（*Three Thousand Years of Urban Growth*）对资料来源注明得较好。对前现代城市规模的所有估计数字都存在争论的空间。钱德勒和福克斯（Fox）的一些估计数字从资料上看是站不住脚的。他们估计的中世纪伊斯兰城市的数据尤其过高，而且像许多历史学家，他们过分夸大了古希腊城市的规模。例如，他们认为公元前430年时雅典

居民有15.5万人，而不是3万~4万人。然而，他们对中古和现代早期中国的估计数字，避免了历史学家们时常出现的浮夸。

只采用单一的资料来源，例如钱德勒和福克斯的《3 000年城市发展史》，然后始终依赖于这个来源，虽然会有一些好处，但似乎弊端更大。依赖于单一资料来源的最大好处是，犯的错误会比较一致，因此容易补偿。然而，对于我们现在讨论的问题，错误似乎会无序散布。所以，我决定不这样做，而是采取无论任何时间、任何地点对专家来说都是最好的办法：交叉检验它们的分数，以减少主观臆断。我总结了西方和东方城市的结果，对于每个城市都列出我的资料来源，以及估计中涉及的任何特殊问题。如果某个估计数字是我自己推测的，我还将提供选择这个数字的理由。我将我对西方的估计收集在表4.1中，对东方的估计收集在表4.2中。研究公元前3000年之前时期的西方考古学家，和研究公元前2000年之前时期的考古学家，以及研究公元第二个千年的历史学家，通常都会提供对城市规模的估计数字，即使他们的数字五花八门。但不幸的是，研究从公元前3000/2000~公元1000年的历史学家和考古学家大多不肯冒险给出具体的估计数字。

从穆赖拜特到纽约

对于每段时间（公元前1400年之前，每100年；公元前1500~前2500年，每250年；公元前2500~前4000年，每500年；公元前5000年之前，每1 000年），我首先提出我确定的最大城市及对其人口的估计，然后提供我的主要资料来源和该城市在社会发展指数中所得的分数，最后简要地评述冲突的估计数字和证据的本质。

表4.1 公元前8000~公元2000年西方最大定居点规模

年代	定居点	规模（人）	分数（分）
公元前8000年	穆赖拜特	大概500	—

公元前7000年	贝达、巴斯塔、恰塔尔休于	1 000	0.01
公元前6000年	恰塔尔休于	3 000	0.03
公元前5000年	布拉克丘	4 000	0.04
公元前4000年	乌鲁克、布拉克丘	5 000	0.05
公元前3500年	乌鲁克、苏萨、布拉克丘	8 000	0.09
公元前3000年	乌鲁克	45 000	0.42
公元前2500年	乌鲁克	50 000	0.47
公元前2250年	阿卡德、孟菲斯	35 000	0.33
公元前2000年	孟菲斯、乌尔	60 000	0.56
公元前1750年	巴比伦	65 000	0.61
公元前1500年	乌鲁克、底比斯	75 000	0.7
公元前1400年	底比斯	80 000	0.75
公元前1300年	底比斯	80 000	0.75
公元前1200年	巴比伦、底比斯	80 000	0.75
公元前1100年	孟菲斯、底比斯、塔尼斯	50 000	0.47
公元前1000年	底比斯	50 000	0.47
公元前900年	底比斯	50 000	0.47
公元前800年	尼姆鲁德/卡尔胡	75 000	0.70
公元前700年	尼尼微	100 000	0.94
公元前600年	巴比伦	125 000	1.17
公元前500年	巴比伦	150 000	1.40
公元前400年	巴比伦	150 000	1.4
公元前300年	巴比伦、亚历山大	150 000	1.4
公元前200年	亚历山大	300 000	2.81
公元前100年	亚历山大，可能还有罗马	400 000	3.75
公元前/公元1年	罗马	1 000 000	9.36
公元100年	罗马	1 000 000	9.36
公元200年	罗马	1 000 000	9.36
公元300年	罗马	800 000	7.49

公元400年	罗马	800 000	7.49
公元500年	君士坦丁堡	450 000	4.23
公元600年	君士坦丁堡	150 000	1.41
公元700年	君士坦丁堡	125 000	1.17
公元800年	巴格达	175 000	1.64
公元900年	科尔多瓦	175 000	1.64
公元1000年	科尔多瓦	200 000	1.87
公元1100年	君士坦丁堡	250 000	2.34
公元1200年	巴格达、开罗、君士坦丁堡	250 000	2.34
公元1300年	开罗	400 000	3.75
公元1400年	开罗	125 000	1.17
公元1500年	开罗	400 000	3.75
公元1600年	君士坦丁堡	400 000	3.75
公元1700年	伦敦、君士坦丁堡	600 000	5.62
公元1800年	伦敦	900 000	8.43
公元1900年	伦敦	6 600 000	61.8
公元2000年	纽约	16 700 000	156.37

公元2000年：纽约，16 700 000人，156.37分。据《经济学人》的《袖珍数字世界》估计，2000年墨西哥城的人口为18 100 000人，圣保罗为18 000 000人，但纽约仍然是西方核心地带（即美国、加拿大边境地区，以及西北欧和中欧）最大的城市。

公元1900年：伦敦，6 600 000人，61.8分。钱德勒估计伦敦的人口为6 480 000人，根据多种官方统计资料，似乎城市历史学家们普遍认同的数字为大约6 500 000。

公元1800年：伦敦，900 000人，8.43分。关于1800年的人口比1900年的争议要多一些。一些资料来源认为伦敦的人口数要小一些，

证据是将官方统计数字与见证人的评述结合在了一起。当时西方第二大城市可能是君士坦丁堡，钱德勒认为其人口为570 000人。

公元1700年：伦敦、君士坦丁堡，600 000人，5.62分。钱德勒估计君士坦丁堡人口为700 000人，而伦敦为550 000人；贝洛赫（Bairoch）认为君士坦丁堡是当时世界上最大的城市，有650 000~1 000 000人。国际中世纪后勤学项目联合主任之一约翰·霍尔顿（John Haldon）认为，君士坦丁堡的人口可能将近700 000人。我的论点是在综合了税收记录、食品进口记录、出生和死亡记录，以及城市覆盖的面积后得出的。

公元1600年：君士坦丁堡，400 000人，3.75分。埃里克·琼斯（Eric Jones）认为君士坦丁堡的人口为600 000人；钱德勒认为是700 000人；贝洛赫则说是650 000~1 000 000人。证据仍然主要是税收记录、食品进口记录、出生和死亡记录，以及城市覆盖面积，但1600年时证据质量骤然下降。

公元1500年：开罗，400 000人，3.75分。弗兰克说贝洛赫估计开罗人口为450 000人，贝洛赫还估计君士坦丁堡人口为300 000~500 000人，但约翰·霍尔顿认为距1453年君士坦丁堡陷落遭洗劫过去不久，人口估计只剩下100 000人左右。证据与1600年和1700年的仍是一类，但对于公元500~1500年，就如何解释这些证据，争议要大得多。欧洲历史学家和中东历史学家有时也采用非常不同的方法，经常导致对伊斯兰城市不切实际的高估，他们认为人口密度会达到每公顷500~1 000人。伊斯兰城市的历史学家对于把握不大的估计往往比欧洲的历史学家谨慎。对开罗的估计似乎尤其成问题。证据主要来自军人登记册、当时人们的印象，以及城市覆盖面积，但在对证据所做的解释中存在很多疑问。

公元1400年：开罗，125 000人，1.17分。这是我本人的估计，是在比较了欧洲城市在黑死病期间极高的死亡率的基础上做出的。钱德

勒认为开罗在1400年仍然有360 000居民，但那将意味着人口只从瘟疫前巅峰时期的450 000人下降了20%。这似乎与阿布·卢格霍德（Abu-Lughod）和多斯（Dols）所述不符。关于证据的本质，见“公元1500年”一段。

公元1300年：开罗，400 000人，3.75分。关于资料来源和问题，见“公元1500年”一段。

公元1200年：巴格达、开罗、君士坦丁堡，250 000人，2.34分。关于这些城市的人口，有一些不同意见，但一般都认为人口全都在200 000~300 000人之间。然而，也有一些估计（尤其是对巴格达），认为还要高很多（见“公元1000年”一段）。

公元1100年：君士坦丁堡，250 000人，2.34分。威克姆（Wickham）认为开罗在11世纪时，人口也达到了250 000人。

公元1000年：科尔多瓦，200 000人，1.87分。这是我的估计。也有几项估计认为科尔多瓦的人口达到了400 000~500 000人。钱德勒也认为君士坦丁堡的人口达到了300 000人，巴格达人口为125 000人。然而，这些估计似乎都太高了。霍尔顿认为君士坦丁堡的人口为150 000人，而巴格达居住区的面积（550~860公顷）似乎太小了，不足以居住100 000人以上的人口。科尔多瓦的面积大致是巴格达的两倍，因此，11世纪其巅峰时的人口为200 000人。

公元900年：科尔多瓦，175 000人，1.64分。这是我的估计。钱德勒估计巴格达的人口为900 000人，君士坦丁堡人口为300 000人，而科尔多瓦为200 000人。还有几位学者把巴格达的人口估得很高，不过再没有人估得像钱德勒那样高了。例如伊拉·拉比达斯（Ira Lapidus）认为是300 000~500 000人，那样的话，人口密度就是每公顷350~900人。按照钱德勒的估计，人口密度则为每公顷1 050~1 600人。这两个

数字似乎都高得出奇了。其他工业革命前的大城市甚至很少有达到每公顷200人的。

公元800年：巴格达，175 000人，1.64分。这是我的估计。巴格达自762年建城后，显然发展非常迅速，在812~813年和865年遭到围攻之前，人口可能达到了巅峰。钱德勒估计的巴格达人口为700 000人，君士坦丁堡为250 000人，科尔多瓦为160 000人。考虑到这些城市的实际面积，以及黑死病肆虐了好几个世纪后西方核心地带普遍人口较少的情况，这些数字似乎又都过高了。霍尔顿估计750年时，君士坦丁堡的人口仅仅为40 000~50 000人。

公元700年：君士坦丁堡，125 000人，1.17分。这是我的估计，是从霍尔顿关于公元500年和750年的数字推测来的。君士坦丁堡的人口在550~750年肯定是陡然下降了。下降自查士丁尼大瘟疫开始，在7世纪10年代的波斯战争和40年代与埃及的谷物贸易瓦解后又加速了。霍尔顿估计750年时，君士坦丁堡的人口为40 000~50 000人，但是证据不足以使我们确定公元700年之前和之后各下降了多少。下降最严重的时期是在公元700年之后。人口在7世纪仅下降了15%~20%，在8世纪又下降了65%。

公元600年：君士坦丁堡，150 000人，1.41人。参见“公元700年”一段的讨论。

公元500年：君士坦丁堡，450 000人，4.23分。卡梅伦（Cameron）和威克姆认为人口应为500 000人，钱德勒认为是400 000人。讨论非常倚重于谷物供应资料。在公元439年丢失了北非之后，罗马人口迅速下降，到公元600年左右时，可能减少到只有20 000~40 000人。威克姆称7世纪的罗马为“都市村庄”。

公元400年：罗马，800 000人，7.49分。罗马人口在公元3世纪时可能就下降了，但很难说下降了多少。不过，很显然在4世纪时，它仍

然是地中海地区遥遥领先的最大城市。到439年，汪达尔人征服北非时，该城也许仍有75万居民。自那以后，人口骤降。威克姆就此提出了一个较低的数字：在5世纪早期时为500 000人。

公元300年：罗马，800 000人，7.49分。参见公元400年一段。公元300年时，城区人口数量比400年时的要低，这也许意味着3世纪时人口的下降比我想象的要剧烈，4世纪时又上升了，但这些都无法肯定。

公元200年，罗马，1 000 000人，9.36分。大多数学者都认为罗马在公元前1世纪晚期时有100万居民，而且大致保持在这个水平上，至少一直到公元200年，然后在3世纪有明显下降，5世纪时急剧下降。不过，我们恐怕已经无法更确切地了解这些情况了。一些学者认为罗马要小得多，人口也许从来没有超过500 000人。然而这有些太小看罗马了，500 000人恐怕是最低的可能数字了。这个论点部分依赖于另一个激烈争论的问题——整个意大利的人口（究竟是四五百万，还是1 200万以上），部分取决于城市本身的人口密度。

公元100年：罗马，1 000 000人，9.36分。这是普遍接受的关于公元头两个世纪的数字（参见“公元200年”一段）。直到大约公元200年之前，人口完全有可能一直在持续增长，但不大可能比100万人多太多。

公元前/公元1年：罗马，1 000 000人，9.36分。参见“公元200年”一段。

公元前100年：亚历山大，也可能是罗马，400 000人，3.75分。谷物贸易记录又一次起到了重要作用。

公元前200年：亚历山大，300 000人，2.81分。

公元前300年：巴比伦、亚历山大，150 000人，1.4分。沙伊德尔认为亚历山大在公元前331年建城后发展非常迅速，但在公元前3~前2世纪速度减缓。

公元前400年：巴比伦，150 000人，1.4分。估计数字是根据城市规模、人口密度，以及希罗多德和亚里士多德等当时的人所做的评述得出的。一些人对巴比伦人口的估计数字要低一些。盖茨（Gates）认为是80 000人，这对于公元前第二个千年的巴比伦是合理的，但对于公元前第一个千年中期的巴比伦，恐怕太低了。

公元前500年：巴比伦，150 000人，1.4分。参见“公元前400年”一段。

公元前600年：巴比伦，125 000人，1.17分。这是我的估计，是根据公元前400年和公元前500年的估计数字推测的。

公元前700年：尼尼微，100 000人，0.94分。估计数字又一次在很大程度上是根据对人口密度的猜测和对诸如《圣经·约拿书》等当时文献的解读得出的。因此，各种估计数字大相径庭。例如，阿凯尔曼（Akerman）认为尼尼微的人口为300 000人，意味着人口密度为每公顷630人。

公元前800年：尼姆鲁德（也被称为“卡尔胡”），75 000人，0.7分。参见“公元前700年”一段。

公元前900年：底比斯，50 000人，0.47分。第三中间期（约公元前1100~前650年）的埃及的文字资料特别匮乏，而考古学家也基本上未对这个时期的遗址发掘予以重视，所以我们对这个时期的估计主要是靠猜测。

公元前1000年：底比斯，50 000人，0.47分。

公元前1100年：孟菲斯、底比斯、塔尼斯，50 000人，0.47分。

公元前1200年：巴比伦、底比斯，80 000人，0.75分。新王国时期的底比斯和青铜时代的巴比伦的居住区如今大都已沉于地下水位之下，使得很难展开正式研究。然而，底比斯显然比中王国时期的城市大得多。中王国的城市一般只有50公顷左右，底比斯也许是公元前1500~前1200年世界上最大的城市了。我们关于巴比伦的信息非常稀少，其中大部分来自德国人早期对默克斯地区的发掘。

公元前1300年：底比斯，80 000人，0.75分。

公元前1400年：底比斯，80 000人，0.75分。

公元前13500年：乌鲁克、底比斯，75 000人，0.7分。还有一些估计要高得多。例如，克里斯蒂安（Christian）认为巴比伦的人口达到了200 000人。

公元前1750年：巴比伦，65 000，0.61分。这是我的估计。我们仍然不知道汉谟拉比（在“长年表”上于公元前1792~前1750年在位）时期巴比伦的面积和人口密度。那时候的巴比伦，如今不仅在地下水位之下，而且深埋在公元前第一个千年的巴比伦之下。它也许是当时世界上最大的城市，统治着一个庞大的帝国。公元前18世纪巴比伦国其他城市的遗迹显示人口密度相当高，因此推测其都城人口为65 000人左右，应当是在正确范围内的，尽管我们缺少能进行严格估计的信息。

公元前2000年：孟菲斯、乌尔，60 000人，0.56分。关于公元前第三个千年的城市（尤其是美索不达米亚的城市）的人口密度，不同的意见实在太多了，以致大多数考古学家都不肯给出具体数字。钱德勒估计的数字相对争议较少。也就是说，我们可以相当有把握地说，公元前第三个千年，甚至是第二个千年，任何城市的人口都不可能达到100 000人，当时最大的一些城市的人口都在50 000人加上15 000人上

下（可得0.33~0.61分）。根据R·M·亚当斯（R. M. Adams）的调查，乌鲁克的人口数字可能比孟菲斯和乌尔还要可靠，尤其比对阿卡德的猜测可靠。目前连阿卡德的位置都还没有确定。

公元前2250年：阿卡德、孟菲斯，35 000人，0.33分。参见“公元前2000年”一段。

公元前2500年：乌鲁克，50 000人，0.47分。参见“公元前2000年”一段。

公元前3000年：乌鲁克，45 000人，0.42分。参见“公元前2000年”一段。

公元前3500年：乌鲁克、苏萨、布拉克丘，8 000人，0.09分。乌鲁克和苏萨的人口数字纯属猜测，而非严格的估计。乌鲁克在公元前3500~前3000年发展得很快。公元前3500年时，该城很显然是最大的苏美尔人定居点，但依照目前已有的证据，我们无法非常准确地推算出其人口。苏萨的遗迹显示该城也是一座大城镇，但由于19世纪的考古发掘质量太差，我们也无法说出该城人口的准确数字。新近在布拉克丘进行的发掘表明，该城在公元前3000年时达到了10 000人，就此前2 000年的水平来说，已算非常大了——甚至可能是当时世界上最大的定居点。然而，至今没有更准确的估计数字了。

公元前4000年：乌鲁克、布拉克丘，5 000人，0.05分。参见“公元前3500年”一段。

公元前5000年：布拉克丘，4 000人，0.04分。参见“公元前3500年”一段。

公元前6000年：恰塔尔休于，3 000人，0.03分。

公元前7000年：贝达、巴斯塔、恰塔尔休于，1 000人，0.01分。耶利哥的规模也许差不多，可能还有一些大致同样规模的更早的定居点；梅塞尔斯（Maisels）认为在公元前8000年左右，穆赖拜特有500~1 000居民。

公元前8000年：最早到公元前7500年前，西方核心地带恐怕没有任何定居点的人数能达到500人，这意味着没有任何城市能达到我所确定的指数的最低分0.01分。

从姜寨到东京

表4.2 公元前4000~公元2000年东方最大定居点规模

年代	定居点	规模（人）	分数（分）
公元前4000年	姜寨、贾湖	300	0
公元前3500年	西坡	2 000	0.02
公元前3000年	大地湾	5 000	0.05
公元前2500年	陶寺、两城镇、尧王城	10 000	0.09
公元前2250年	陶寺、两城镇、尧王城	14 000	0.13
公元前2000年	鄆城南水	11 000	0.10
公元前1750年	二里头	24 000	0.22
公元前1500年	郑州	35 000	0.33
公元前1400年	郑州	35 000	0.33
公元前1300年	郑州	35 000	0.33
公元前1200年	安阳	50 000	0.47
公元前1100年	安阳	50 000	0.47
公元前1000年	洛邑、酆京	35 000	0.33
公元前900年	洛邑、酆京	40 000	0.37
公元前800年	洛邑、酆京	45 000	0.42
公元前700年	临淄、洛邑	55 000	0.51

公元前600年	临淄、洛邑	65 000	0.61
公元前500年	临淄	80 000	0.75
公元前400年	临淄、曲阜、洛邑、新郑、舞阳	100 000	0.94
公元前300年	临淄、曲阜、洛邑、新郑、舞阳	125 000	1.17
公元前200年	长安	250 000	2.81
公元前100年	长安	375 000	3.75
公元前/公元1年	长安	500 000	4.68
公元100年	洛阳	420 000	3.93
公元200年	长安	120 000	1.12
公元300年	平阳、长安、洛阳、许昌、邺	140 000	1.31
公元400年	平城	200 000	1.87
公元500年	洛阳	200 000	1.87
公元600年	大兴城/长安	600 000	5.63
公元700年	长安	1 000 000	9.36
公元800年	长安	1 000 000	9.36
公元900年	长安	750 000	7.00
公元1000年	开封	1 000 000	9.36
公元1100年	开封	1 000 000	9.36
公元1200年	杭州	1 000 000	9.36
公元1300年	杭州	800 000	7.50
公元1400年	南京	500 000	4.68
公元1500年	北京	678 000	6.35
公元1600年	北京	700 000	6.55
公元1700年	北京	650 000	6.09
公元1800年	北京	1 100 000	10.3
公元1900年	东京	1 750 000	16.39
公元2000年	东京	26 400 000	250.00

公元2000年：东京，26 400 000人，250分。中国最大的城市为上海，12 900 000人，120.79分。

公元1900年：东京，1 750 000人，16.39分。一些城市历史学家做出了稍低的估计，但是对这一地区的估计数字，意见总体似乎是一致的。估计数字是根据人口普查、纳税申报、食品供应和军事人员等多种官方统计资料做出的。当时中国最大的城市是北京，大约有1 100 000居民（10.3分）。

公元1800年：北京，1 100 000人，10.3分。对清朝时期北京的估计，很大程度上依据于调运粮食的统计资料，估计结果大相径庭。在不同的时间点上，布罗代尔（Braudel）认为北京的人口为300万或200~300万。钱德勒的估计数字似乎与社会历史学家们对清代北京的描述更为相符。

公元1700年：北京，650 000人，6.09分。在遭受1644年残酷的洗劫之后，北京人口急剧下降，到1700年时可能仍未恢复到1600年的水平。然而，也有一些历史学家给出了较高的人口数字。

公元1600年：北京，700 000人，6.55分。一些历史学家给出了更高的数字，但却没有提供支持这些数字的证据。

公元1500年：北京，678 000人，6.35分。莫特（Mote）估计16~17世纪南京和北京的人口都在大约100万左右，但这似乎不大可能，既因为这个数字太高了（北京可能直到18世纪晚期，人口才达到100万），也因为学者们普遍认为自1421年北京取代南京成为国都以来，南京人口下降了大约50%，这是莫特在别处也承认了的。贝洛赫也同意较低的估计数字，他认为1600年时北京至少有600 000人。

公元1400年：南京，500 000人，4.68分。莫特说他认为南京的人口大约有100万人，但根据他本人粗略的计算产生的结果实际上是400

000~500 000人。

公元1300年：杭州，800 000人，7.5分。贝洛赫认为1300年左右另有4个中国城市人口在200 000~500 000人，而杭州“可能要大得多”。然而，他根据稻米的消费量计算而来的数字更准确地指向了800 000人，同时，伊懋可通过稻米消费量计算出的数字为600 000~700 000人。饶济凡（Rozman）也认为12~13世纪杭州的人口在500 000人以上，甚至可能高达100万人。库恩（Kuhn）和克里斯蒂安也倾向于100万人，施坚雅（Skinner）认为是120万人。我为1200年选择较高的数字：大约100万；为1300年选择较低的数字：800 000人。到1300年时，中国整体人口都减少了。当马可·波罗于13世纪晚期访问杭州时，该城无疑是当时世界上最大的城市，但是马可·波罗言语间暗示的数字（500万~700万）肯定是高得离谱了。马可·波罗恐怕根本不可能知道杭州的人口数，只需想想该城比他那个时代欧洲和伊斯兰世界的城市都大得多即可。

公元1200年：杭州，1 000 000人，9.36分。参见“公元1300年”一段。

公元1100年：开封，1 000 000人，9.36分。钱德勒和贝洛赫都认为开封较小（他俩分别认为开封的人口为400 000人和400 000~450 000人），但这似乎与当时文献对该城市的描述不符。不确定因素似乎很大程度上来自哪些区域应算作“城市”这个问题。新城建于955年，有长27千米的城墙（962年又延长了3.3千米），在旧城46个行政区的基础上又增加了75个行政区，但在公元1000年之前，该城人口就已经在向城墙之外发展了。到1021年时，14个新的城外大行政区得到了承认。官方统计资料称，980年左右开封辖区共有890 000人定居，1103年时增加到130万人，城市某些部分的人口密度达到了每公顷500人。如果我们只计算城墙内的人口，那么钱德勒和贝洛赫的估计可能是合理的；如果我们计算全部人口，莫特、施坚雅和库恩选择的官方数字，

看来就是合理的了。我倾向于后者，但考虑到资料的模糊性，我索性做一个粗略的估计：100万人。按照官方数字，1100年时，杭州可能有80万~100万人。

公元1000年：开封，1 000 000人，9.36分。参见“公元1100年”一段。

公元900年：长安，750 000人，（我的估计）7分。中国的历史学家很少对公元900年左右的长安人口发表意见。9世纪70年代末，黄巢曾反复洗劫过这座城市，并于880年和883年将其彻底焚毁，其人口锐减也就毫不奇怪了。在9世纪70年代末之前，长安无疑是世界上最大的城市。贝恩（Benn）认为其人口达到了200万人，库恩认为“在100万人以上”，但即使已经开凿了大运河，也很难看出怎么可能运输那么多谷物到长安，来供养贝恩提出的那么多人口。施坚雅认为中唐时期长安可能有大约100万人，似乎更为可信，我认为这个数字适用于公元800年和700年。包围了30多平方英里[注]的城墙内无疑能容下100万人，但200万人恐怕不大可能，人口密度实在太大了。不过，自9世纪70年代以来，人口是怎样骤降的，也不大清楚。原始资料称当皇帝唐僖宗于885年回到长安时，该城已被彻底毁灭，但这显然是夸大之词，因为该王朝又在那里维系了20年，直到904年军阀朱温下令将所有仍存在的建筑摧毁。直到那一年之前，长安仍是主要的人口中心。不过，即使这个猜想是错误的，公元900年时东方的社会组织/城市规模的分数仍然很高，因为洛阳那时可能也有500 000~750 000人。武则天于7世纪晚期移居洛阳时，可能有100 000多户家庭随之迁移。贝恩认为洛阳的人口高达100万人。不过，饶济凡提出的数字是500 000人。

公元800年：长安，1 000 000人，9.36分。参见“公元900年”一段。

公元700年：长安，1 000 000人，9.36分。参见“公元900年”一段。

公元600年：大兴城（7世纪时被唐朝重新命名为“长安”），600 000人，5.63分。这是我的估计。隋朝兴建了大兴城，作为其新都，城墙包围的面积有30多平方英里，在7世纪时可容纳约100万人。不过，皇帝于583年定都于此时，该城还仍在建设中，很多行政区尚无人居住。公元600年时，该城人口可能已经很多了，因为建筑工程需要数以万计的劳工，加上他们的家眷，更不用说大量官员和工匠（及其家眷），还有上百座寺庙里数以千计的和尚和尼姑了。而且，当隋朝于589年灭亡了中国南方的陈朝后，又有大量人口从南方迁到了大兴城。

公元500年：洛阳，200 000人，1.87分。北魏的孝文帝于493年将都城从平城迁到了洛阳。据史书记载，他还于495年调集了150 000名士兵到洛阳，并将部分洛阳周边土地赐予他们。该城于公元6世纪扩展了许多，人口很可能像大兴城一样，达到了600 000人。

公元400年：平城，200 000人（我的估计），1.87分。公元400年左右，中国北方有好几座大城市，但平城（1048年更名为“大同”）可能是最大的。史书记载公元398年，有100 000名鲜卑人被强制迁往平城；399年时，又有100 000名河南农民和2 000户富裕的中国少数民族家庭被迁往那里。除了像邺城等极少数例外之外，公元200~400年这段时期城市的考古证据极其匮乏。

公元300年：平阳、长安、洛阳、许昌、邺城，140 000人（我的估计），1.31分。公元四五世纪时，很难定义究竟什么样才能算城市；因为那个时期的主要战争基本上都是来自奴隶的袭击，所以在中国北方，城市就像是巨大的军营。军阀聚拢了成千上万户家庭，将他们围在中央，而堡垒周围则有广袤的可供耕种的土地。平阳、长安、洛阳、许昌和邺城在公元300年左右都成了大城市，可能比公元200年左右最大的城市还要大，而比公元400年左右最大的城市稍小一些。

公元200年：长安，120 000人（我的估计），1.12分。公元190年时，军阀董卓劫掠并摧毁了洛阳，将其人口迁移到长安；196年，曹操将皇室也迁到了长安（直到他死后，皇室才迁回洛阳）。这些城市显然比公元100年的洛阳要小得多，更不用说公元1年的长安了。

公元100年：洛阳，420 000人，3.93分。考古学家和历史学家非常详细地描绘了汉朝一些主要城市的布局，却没有提供对人口的估计。从考古发掘报告和留存下来的城市平面图来看，似乎长安和洛阳（在公元前206~公元32年和公元32~220年两段时期的大部分时间，分别为都城）的人口都有数十万。据史书记载，秦始皇曾于公元前3世纪20年代强行迁移了120 000户家庭到其都城咸阳，在公元前3世纪10年代又迁移了更多家庭去照料他的墓地。这些数字很可能有所夸张，但在公元前210年秦始皇驾崩时，咸阳可能的确有200 000人以上的居民。汉朝在长安的新都城至少也一样大。到公元前1世纪时，长安的两个主要市场面积分别达到了50公顷和25公顷，这也说明了人口之庞大。该城面积广大，达到了44.5平方公里，但从发掘区域的人口密度，以及一向不够完备的长安粮食供应记录来看，其人口数量不及当时的罗马。我估计到西汉末年时（即公元前/公元1年左右），该城的人口可能达到了巅峰，为大约500 000人，不过这种猜测的误差很可能会达到20%左右。

长安周围还有一些卫星城，使得估计其人口变得更为复杂。尤其是围绕着皇陵发展起来的卫星城，分别散布在郑国渠两岸30千米和灞河、浐河两岸20千米的区域内。如果我们把长安及其卫星城合并在一起计算，那总人口会超过罗马，但由于这些卫星城从各方面来看都像是独立的城市，所以我没有这样做。也有一些证据表明长安在公元前100年后发展放缓，自汉武帝于公元前87年死后，也没有再建立新的卫星城。洛阳面积比长安小一些，但人口显然比长安稠密。我因此对洛阳做稍低的估计，认为其人口在公元100年巅峰时为420 000人。当然，如有上下20%的误差也是合理的。

公元前/公元1年：长安，500 000人（我的估计），4.68分。参见“公元100年”一段。

公元前100年：长安，375 000人（我的估计），3.75分。参见“公元100年”一段。

公元前200年：长安，250 000人（我的估计），2.81分。参见“公元100年”一段。

公元前300年：临淄、曲阜、洛邑、新郑、舞阳，125 000人（我的估计），1.17分。关于春秋战国时期城市的考古资料依然很少，但这些城市的规模在公元前第一个千年的下半叶似乎明显稳步扩大。最大的城市的城墙[舞阳（燕国），长达27千米；新郑（郑国/韩国），16千米；临淄（齐国），15千米；曲阜（鲁国），14千米；洛邑，后更名为洛阳（周朝都城），12千米]一般都围拢了9~15平方千米的面积，说明人口应在100 000~200 000人之间。然而，一些城市显然有较大的典礼和工业用地，而且（至少在起初）可能有大片的土地被并入城墙内，只是为未来的发展预留。下面的估计是我自己做出的，误差可能会比汉朝的城市大，甚至可能高达上下50%。

古代的文献资料并无多大帮助。《史记》中记载齐国的临淄有70 000户人家，号称有210 000名成年男子。司马迁评论说，该城如此拥挤，以致人们“挥汗如雨”。他的数字暗示总人口可能达到350 000~750 000人，这将使临淄的人口比同时代的巴比伦要多得多。不过，考虑到该城的实际面积，似乎是不大可能的；而且这还意味着中国人口最多的一些城市在公元前500~前1年之间实际上没有增长，即使证据毫不含糊地表明它们的面积在这段时期至少增加了一倍，甚至可能增加了4倍。

贝洛赫认为战国时期（公元前480~221年）有4~6座城市的人口在100 000人以上，这与我此处所做的估计非常一致。

公元前400年：临淄、曲阜、洛邑、新郑、舞阳，100 000人（我的估计），0.94分。参见“公元前300年”一段。

公元前500年：临淄，80 000人（我的估计），0.75分。参见“公元前300年”一段。

公元前600年：临淄、洛邑，65 000人（我的估计），0.61分。公元前第一个千年上半叶的考古证据，甚至比下半叶还要缺乏（其实比公元前第二个千年晚期都要少）。我们可以肯定公元前1000年左右最大的城市，比公元前500年左右的要小，但我们不能确定小多少。我猜测前者的人口大约是后者的一半，但是这全都依赖于对定居点面积和人口密度的估计。

关于最大城市（西周的都城、渭河流域的酆京和镐京，东周的都城洛邑）的资料非常少，而且大多局限于贵族陵墓和青铜器皿。酆京的出土物散布在大约12.5平方千米的区域内，而镐京的出土物散布在大约6平方千米的区域内，但这些区域只有小部分当年是盖有建筑物的。对于洛邑，我们甚至不知道那些偶然发现的出土物就是来自洛邑城本身，还是也有可能来自郑州。

冯·法肯豪森（Von Falkenhausen）认为“周原（即酆京和镐京地区）上的西周都城，是由一些宗教兼居住功能的大院落非常随意的聚合而构成的。这些院落散布在可能广阔达200平方千米的区域内。院落之间有大片的农田相隔”。如果他的说法是正确的，那就不仅意味着周朝都城人口非常之少，而且意味着由于居住模式过于分散，公元前第一个千年早期的中国根本谈不上“城市”。这种情况也适用于公元前第二个千年晚期的“城市”。

即便如此，这200平方千米内的出土物显然密度不同，（像公元前第一个千年的中国作家们那样）认为酆京、镐京和洛邑是明显的核心地带，似乎是合理的，即使从有密集、连续的房屋建筑区域这个意义

上判断它们不是严格的“城市”。我猜测公元前1000年左右的洛邑和酆京有35 000人，镐京可能有它们的一半。我想公元前1000年时洛邑和酆京还不大可能有50 000人，因为公元前第一个千年上半期的增长量不可能超过20 000人。因此，东方最大的城市是以相当平稳的增长率发展的，从公元前1000~公元500年，人口增长了一倍稍多，从35 000人增长至约80 000人。

公元前700年：临淄、洛邑，55 000人（我的估计），0.51分。参见“公元前600年”一段。

公元前800年：洛邑、酆京，45 000人（我的估计），0.42分。参见“公元前600年”一段。

公元前900年：洛邑、酆京，40 000人（我的估计），0.37分。参见“公元前600年”一段。

公元前1000年：洛邑、酆京，35 000人（我的估计），0.33分。参见“公元前600年”一段。钱德勒认为洛邑人口有50 000人。

公元前1100年：安阳，50 000人（我的估计），0.47分。作为商朝最后的都城，安阳自1928年以来得到了大规模发掘，尽管有城墙的城市洹北直到1997年才被定位。洹北城墙包围的面积有470公顷，判断其中有20 000~25 000人似乎是合理的，但安阳的其他遗迹蔓延了大约30平方千米。像公元前第一个千年早期一样（参见“公元前600年”一段），在这样一个分散的定居系统中，很难确定一个“城市”的边界在哪里。因此我提出的50 000人就多少有些武断了。如果非常狭义地将这个城市只定义为城墙包围的区域，那么这个估计数字可以砍去50%；而如果非常广义地将其定义为包含郊区，则又可能将总人口提高到100 000人，甚至还要多。如果安阳的人口有50 000人，就与公元前1100年的孟菲斯一样大了；如果有100 000人，那它就是公元前13~

前11世纪世界上最大的城市了。我提出50 000这个数字，是对这座城市非常狭义的定义和非常广义的定义取中。

安阳建于约公元前1300年，到公元前1200年时已显然成为主要的定居点（无论怎样定义）。考虑到对公元前1100年做估计时的不确定性，为公元前1200年提出一个不同的估计数字，将更麻烦也没有什么意义，所以我索性为两个年代都提出50 000人这个数字。

公元前1200年：安阳，50 000人（我的估计），0.47分。参见"公元前1100年"一段。三星堆有围墙的定居区面积也许能达到350公顷，人口也许能与安阳匹敌，但我们对三星堆仍然知之甚少。

公元前1300年：郑州，35 000人（我的估计），0.33分。郑州的二里岗遗址，大约兴建于公元前1600年，通常被认为是商朝早期的一个都城。有围墙的定居区面积有300公顷，但还有一堵更大的外围墙包围了总计达1300公顷的区域。像安阳（参见"公元前1100年"一段）一样，想准确估计其人口有两大难点：第一，在这样一种情况下如何定义"城市"；第二，如何，计算城市内的人口密度。我只好又一次在最狭义的定义（城墙之内的核心区域人口不到15 000人）和极广义的定义（多达50 000人）之间取中了。郑州似乎比公元前13~前11世纪的安阳要小得多；我估计的数字为35 000，将使其规模相当于同时代的巴比伦和底比斯的一半。

公元前1400年：郑州，35 000人（我的估计），0.33分。参见"公元前1300年"一段。在缺乏详细的考古证据的情况下，我对公元前16~前14世纪的郑州提出同样的数字。

公元前1500年：郑州，35 000人（我的估计），0.33分。参见"公元前1400年"和"公元前1300年"两段。

公元前1750年：二里头，24 000人，0.22分。二里头比公元前1500~前500年的遗址得到了好得多的探索。二里头三期的面积大致达到了300公顷。
24 000人这个估计数字——即使刘莉认为只是作为从18 000~30 000人的一系列估计数字的取中点提出的——可能仍然是东方最可靠的史前人口统计学数据了。24 000人这个数据意味着大约每公顷80人，按照巴比伦等同时代的西方城市的标准来看，人口密度算是低的，但与史前其他中国定居点比，仍相对较高。

公元前2000年：鄩城南水，11 000人（我的估计），0.1分。这个定居点面积似乎达到了230公顷，但发掘得仍然非常不足。我猜测人口密度人较低，大约每公顷50人。

公元前2250年：陶寺、两城镇、尧王城，14 000人（我的估计），0.13分。陶寺在其全盛时期，面积达到280公顷，我猜测人口密度为每公顷50人。刘莉也说龙山文化时期最大的酋邦可能有10 000多名成员，这也许意味着对陶寺的人口密度，我们应当采用较低的数字（甚至即使以史前中国定居点的标准来看，陶寺的出土物实际上极其分散）。新近的研究表明，两城镇和尧王城在公元前第三个千年的下半期，面积甚至比陶寺还要大，分别达到了272.5公顷和367.5公顷。

公元前2500年：陶寺、两城镇、尧王城，10 000人，0.09分。公元前2500年的陶寺显然比其后来的巅峰时期要小，但我不知道有什么更好的估计数字。参见“公元前2250年”一段。

公元前3000年：大地湾，5 000人（我的估计），0.05分。该定居点的面积大约为100公顷，我猜测人口密度在每公顷50人左右。

公元前3500年：西坡，2 000人（我的估计），0.02分。该定居点的面积大约为40公顷，我猜测人口密度在每公顷50人左右。

公元前4000年：没有任何定居点面积大到能供养1 000人，故而得分为最低分（0.01分）。公元前4000年时，姜寨的面积为5公顷，刘莉计算的人口密度为每公顷44~63人，意味着该定居点只有220~315人。彼得森（Peterson）和谢拉赫（Shelach）为该遗址的人口设计了一个有趣的动态模型，所产生的数字要稍高一些，但也仅仅是400人左右。公元前6000年时，贾湖的面积也是大约5公顷，人口密度也非常低。从公元前第7个至公元前第5个千年，似乎再没有面积能超过2公顷的考古遗址了。

多少人才够

间接度量

自我们有文献资料起（西方始于公元前第三个千年，东方始于公元前第二个千年晚期），直到公元20世纪，在所有时间点上，世界上最大的城市都是行政中心。从有文字记载的历史开端起，孟菲斯是埃及的都城，而安阳是商朝的都城；到了19世纪，伦敦是大英帝国的首都，而北京是大清帝国的首都。如果我们将时间回溯到孟菲斯和安阳之前，有一定数量的证据表明，西方的乌鲁克和东方的郑州（也可能是二里头），也是更早期国家的都城。

这个现象似乎证明了选择城市规模作为社会组织的间接度量手段的正确性。在大部分时间里，一个地区最大城市的规模就是衡量其政治组织职能的尺度之一。在以前发表的一篇论文中，我曾经提出公元前第一个千年希腊世界的情况正是如此，现在我要将这一论点扩展到整个前现代历史。只是到了20世纪，经济力量才压倒政治力量，以致世界上最强大的国家的首都华盛顿在2000年未能跻身于世界最大城市的前30名；而东方最强大国家的首都北京，只排在第24位。在此前的全部历史中，城市规模都是社会组织的非常直接的反映。

能量获取失去对城市规模的影响

总体而言，能量获取历史的图形（见图2.5）与城市规模–社会组织的图形（见图4.1），在一定程度上是相同的。两者都在冰期结束后增长非常缓慢，在公元前最后几千年开始加速，然后在公元19世纪和20世纪呈爆炸式增长。在两幅图中，在过去10 000年间，西方的分数在大部分时间都高于东方。然而，两幅图中的差异也如相同处一样引人注目。

图4.1 公元前8000~公元2000年东方和西方最大的城市的规模

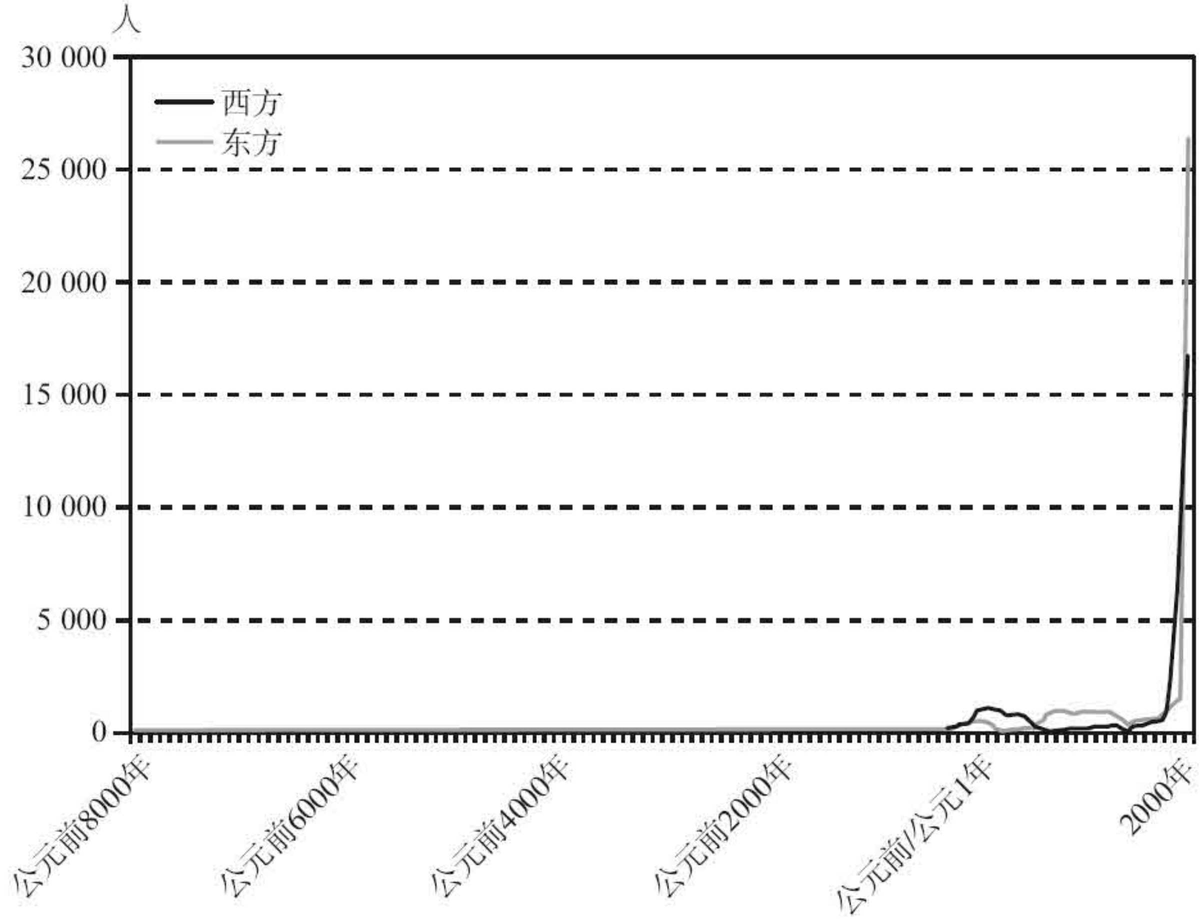


图4.2和图4.3分别标绘了西方和东方的能量获取情况和城市规模（以社会发展指数上的分数的形式表示），以及两者在对数–线性标尺上的对比（图4.4和图4.5则在线性－线性标尺上展示了同样的数据；可

以看出同样的模式，不过对比不像对数标尺上那样鲜明）。能量获取曲线与城市规模曲线最引人注目的反差，似乎是：（1）城市规模曲线开始上升的时间比能量获取曲线要晚得多；（2）城市规模曲线比能量获取曲线不稳定得多。这两个反差很容易解释：城市规模是能量获取的功能之一。只有当能够达到某种水平的能量获取（每人每天7 000~8 000千卡），最大定居点的规模才会开始显著增长；然而一旦某个社会跨过了这个门槛，能量获取预算边缘相对较小的变化，就能对可用于组织较大的社会的能量数量产生巨大影响。

图4.2 公元前14000~公元2000年西方能量获取情况与城市规模在对数–线性标尺上的对比（以社会发展分数衡量）

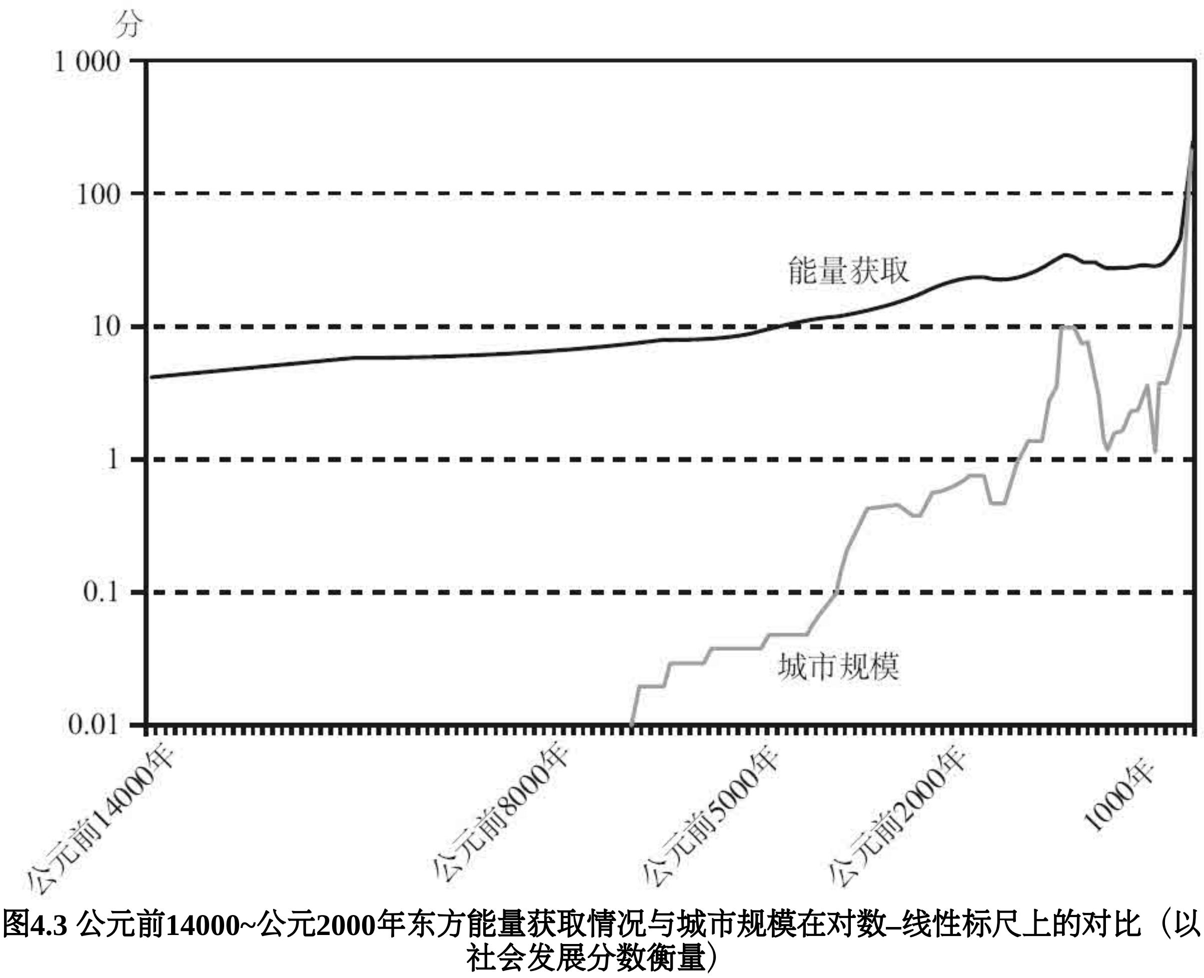


图4.3 公元前14000~公元2000年东方能量获取情况与城市规模在对数–线性标尺上的对比（以社会发展分数衡量）

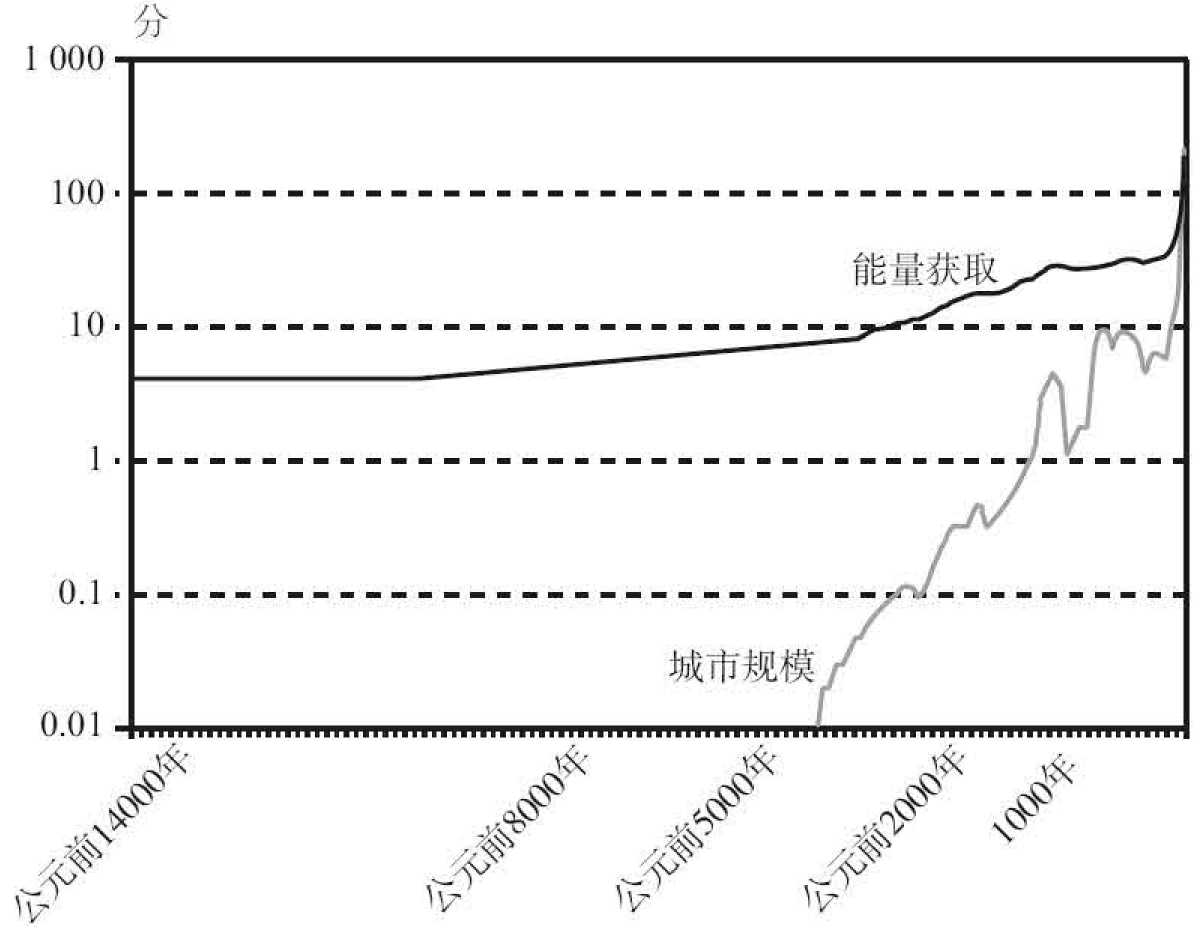


图4.4 公元前14000~公元2000年西方能量获取情况与城市规模在线性－线性标尺上的对比（以社会发展分数衡量）

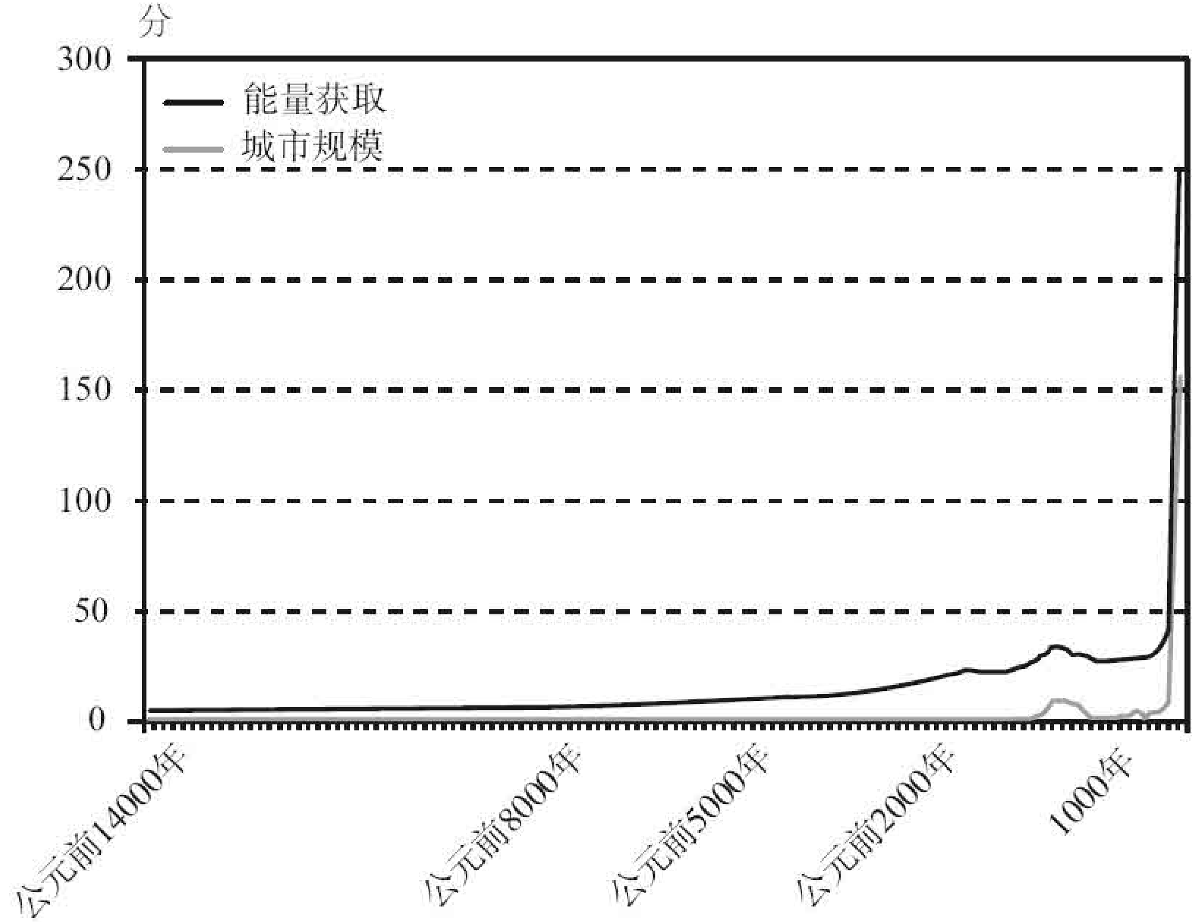


图4.5 公元前14000~公元2000年东方能量获取情况与城市规模在线性－线性标尺上的对比（以社会发展分数衡量）

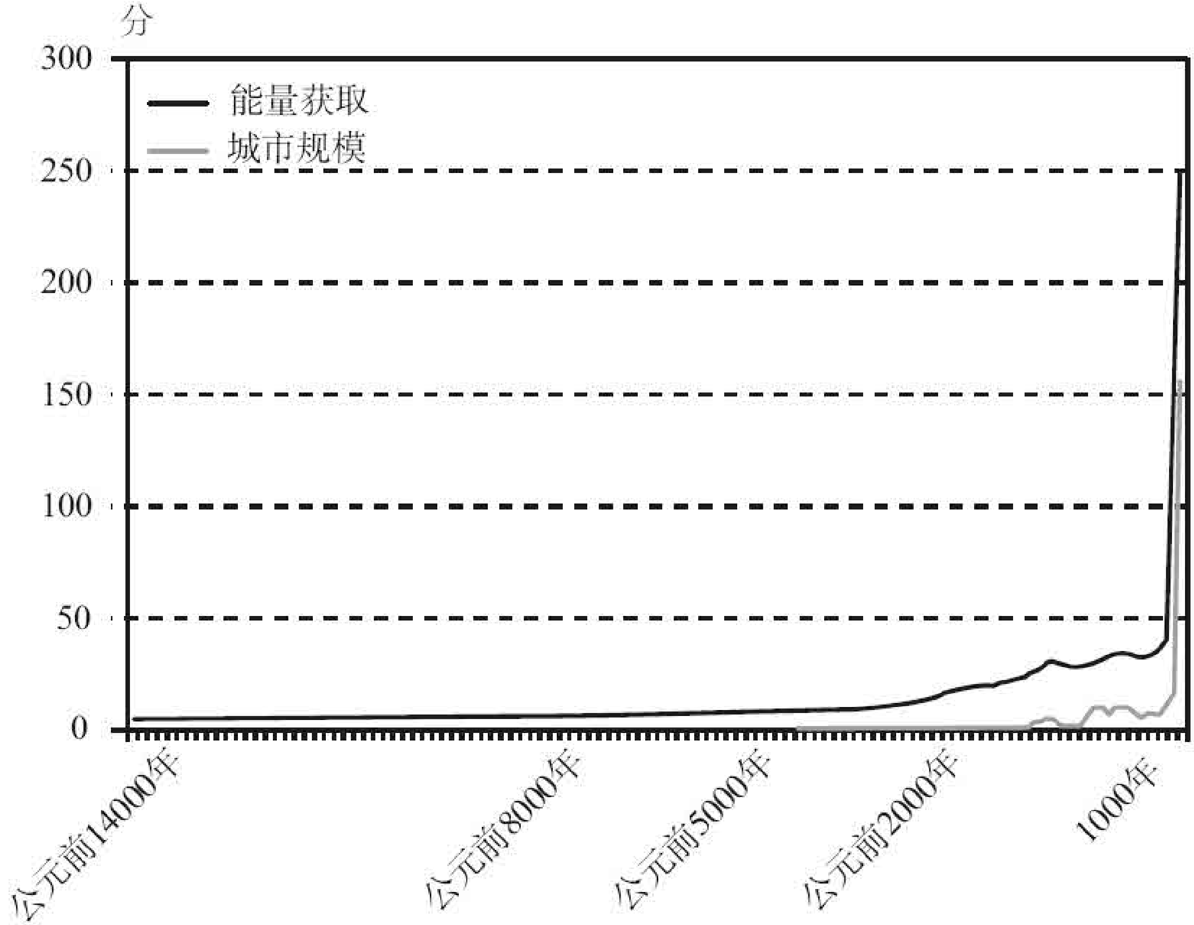
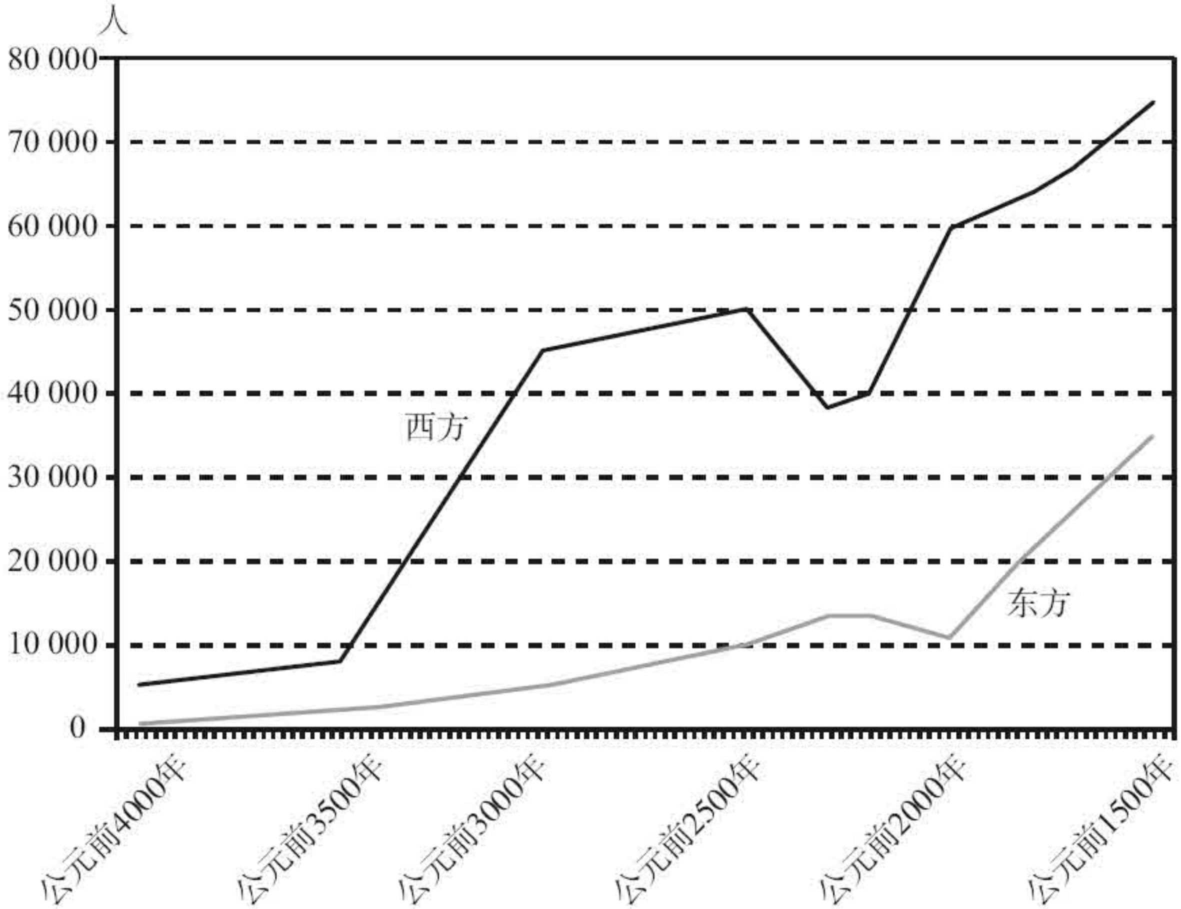


图4.6 公元前4000~前1500年东方和西方最大定居点规模

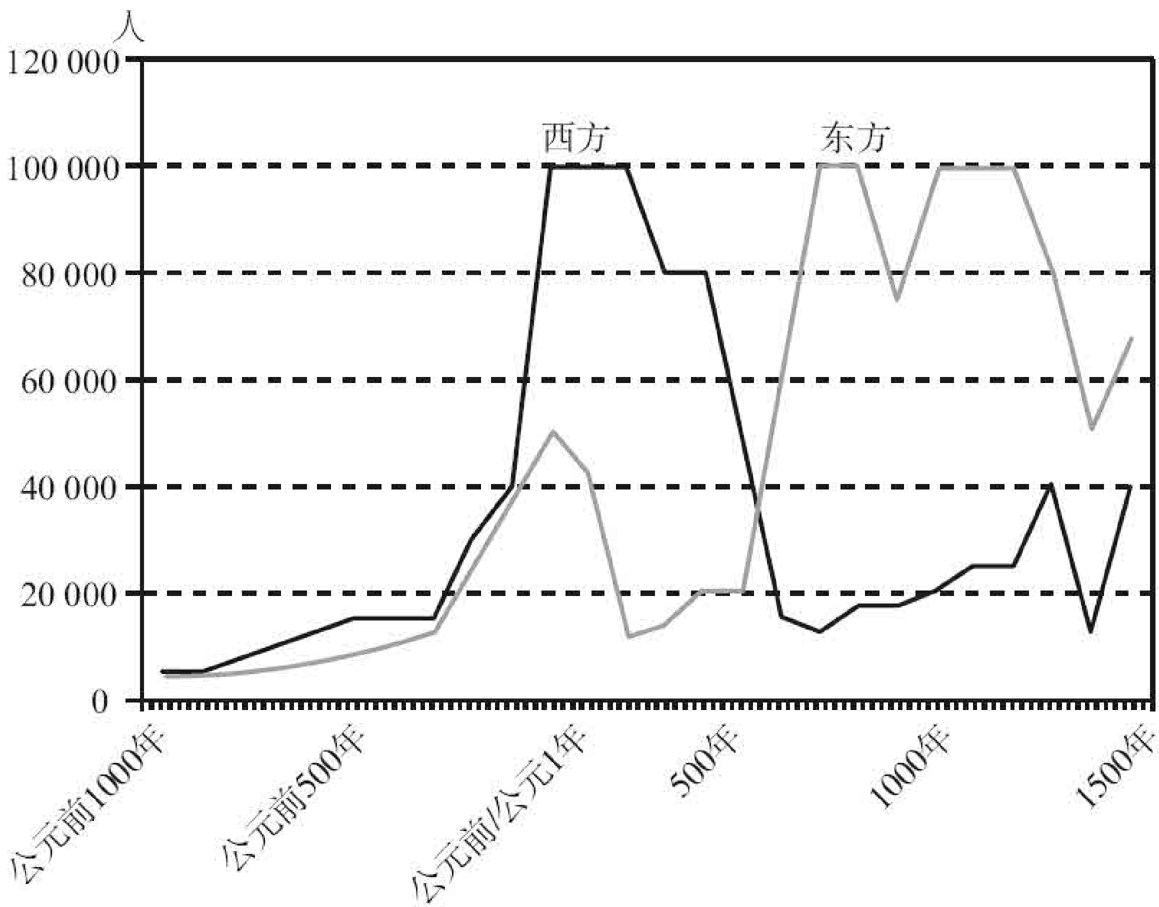


因此，无论在东方还是在西方，当能量获取水平达到每人每天大约11 000~12 000千卡时（见图4.6；西方于大约公元前3500~前3000年，东方于公元前2000~前1500年到达此水平），城市化的起步阶段历程很相似。在公元前第三个千年末期时，两者的定居点规模都出现了暴跌。西方发生了阿卡德、乌尔陷落和埃及旧王国衰败等危机，东方则出现了陶寺和山东早期城市的败落，尽管这些危机对东西方的能量获取状况都只产生了极微小的影响。

最近3 000年的变化甚至更为惊人（见图4.7）。无论在东方还是在西方，能量获取的增长率在公元前第一个千年都加速提高了，但城市规模的增长还要更快。能量获取似乎又出现了一个门槛，这回是略高于每人每天20 000千卡。跨过了这道门槛的社会，就能创造出有100 000人以上的城市。还有一道门槛是大约每人每天27 000千卡，跨过这

道门槛，创造有50万~100万人口的超级城市便有了可能。公元第一个千年早期的大危机在东方和西方都造成了能量获取能力前所未有的锐减（在西方，公元100~700年降低了将近20%；在东方，公元100~300年降低了将近4%），但危机对城市规模的影响还要大得多：公元200~700年，西方的城市萎缩了85%以上；公元1~200年，东方的城市萎缩了75%以上。

图4.7 公元前1000~公元1500年东方和西方最大定居点的规模



公元第一个千年的中晚期，东方的城市规模急剧扩大。在其跨过了每人每天27 000千卡的能量获取门槛后，出现了堪与公元前第一个千年晚期的罗马相匹敌的城市。东方的能量获取水平在公元500~1000年增长了13%（从每人每天26 000千卡增至29 500千卡），但是东方的城市规模同样在这500年间增长了400%（从20万居民增长至100万居

民）。公元第一个千年晚期颠覆了中国唐朝的战争几乎没有对能量获取水平产生什么影响，但的确在短期内造成城市规模下降了25%。

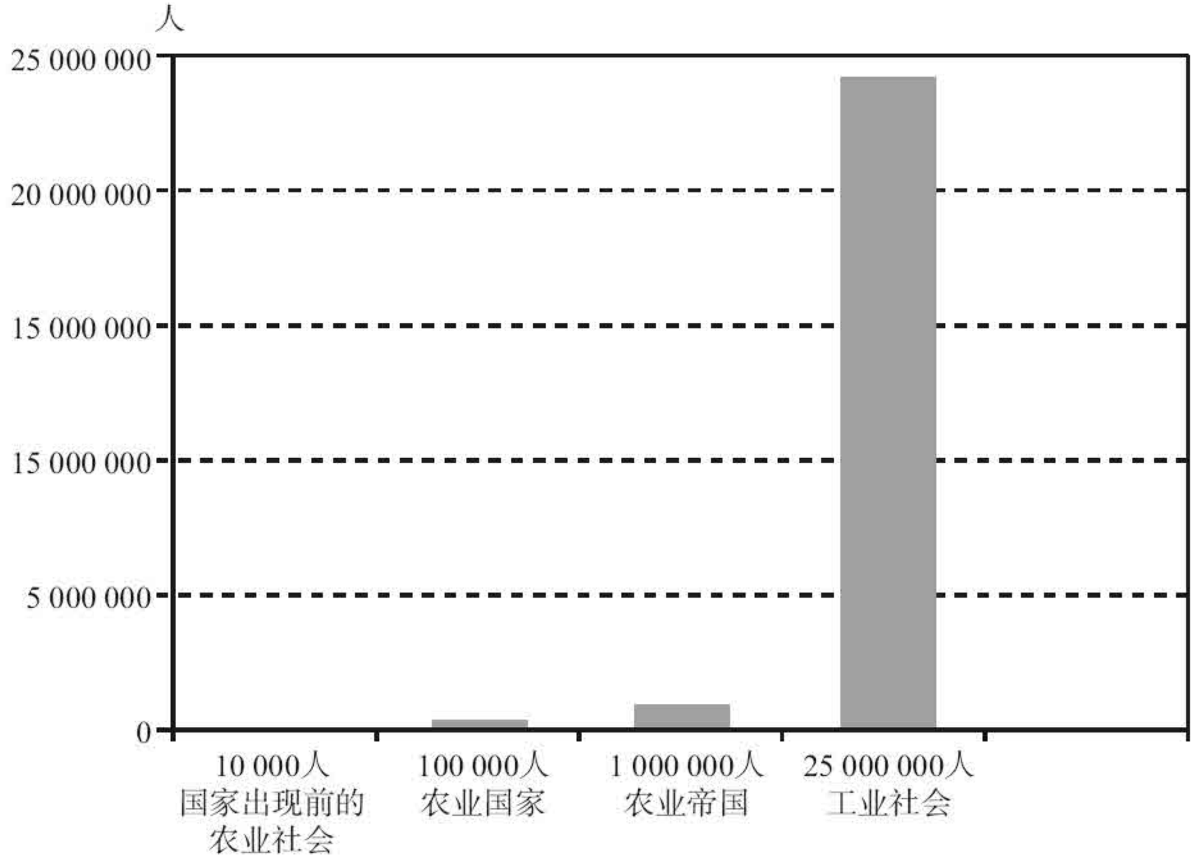
能量获取与城市规模的关系一直到了公元第二个千年都还在起作用。1200~1400年旧世界的第二次大变迁导致东方的能量获取水平下降了5%，但最大城市的人口下降了一半；在西方，能量获取水平未受影响，但城市规模萎缩了几乎2/3。

能量获取水平自公元1500年（特别是1800年后）的提升，对城市规模有可想而知的巨大影响。大约每人每天45 000千卡似乎是能量获取的又一道门槛，使得出现数以百万计居民的城市成为可能。20世纪的世界大战使得东方最大的一些城市受到严重破坏，但城市规模的波动性已经极强，东京和北京在20世纪60年代以后都发展得比以往更大，而西方最大的一些城市（在美国）在两次大战中均毫发无损。

城市规模的量级

城市规模数据也表明，不同的社会发展水平决定着定居点规模的量级。国家出现以前的农业社会（如在西方核心地带公元前3500年之前和东方核心地带公元前2000年之前出现的社会）似乎无法供养大约10 000人以上的定居点；农业国家（公元前第4个到前第一个千年早期主宰着西方核心地带，公元前第二个千年早期到公元前第一个千年中期主宰着东方核心地带）似乎供养不起大约10万以上人口的定居点；农业帝国（公元前第一个千年中期到公元第二个千年晚期主宰着西方核心地带，公元前第一个千年晚期到公元第二个千年晚期主宰着东方核心地带）似乎供养不起大约100万以上人口的定居点。然而，工业社会却能够维系2 500万以上人口的城市（见图4.8）。

图4.8 自冰期以来已知最大定居点及群落组织水平



前现代时期有序的层层递进，当然部分是因为定量估计的粗糙性（图4.7中线顶点是平的纯粹表明我们资料的缺乏；罗马、长安、开封和杭州拥有80万或120万人口，与图中标绘的它们拥有100万人口，是同样有可能的）。然而，结果的一致性的确表明了一个值得以世界上其他地方的数据进行检验的假设——如果没有化石燃料带来的能量激增，以及与之相关的组织和技术方面的收益，则没有人能生活在人口超过100万的城市中。我们还需要观察，我们目前的发展水平会给城市规模设置怎样的上限，以及我们是否能突破这个上限。

---

1. 1英里≈1609米；1平方英里≈258万平方米。——编者注

The Measure of Civilization

How Social Development Decides the Fate of Nations

第五章

战争能力：工业变革带来的巨大差距

量化战争能力

再没有什么事情像1840~1842年的第一次鸦片战争那样，使得西方主宰世界的局面变得如此清楚。一支小小的英国舰队在中国长驱直入，威胁要切断向北京运送粮食的大运河，并迫使清政府做出了丧权辱国的让步。根据随军的罗伯特·乔斯林（Robert Jocelyn）勋爵的记录：“军舰向定海小城舷炮齐发，房倒屋塌、木头相撞、人们呻吟哀号的声音在岸上回荡。我方的炮击持续了9分钟……我们在一处废弃的河滩上上了岸，除了几具尸体，一些弓和箭、折断的矛和枪，这片地区什么都没有了。”

中国人很好地接受了教训。“所有共产党人都必须明白这个道理。”毛泽东在一个世纪后说：“枪杆子里面出政权。”事实一向如此，战争能力始终是社会发展的关键因素。

对社会发展指数来说，幸运的是，许多因素——历史学家对记述战争的痴迷，强制性军事档案的保存，艺术赞助者们喜好被描绘成武士，普遍存在的以武器和盔甲陪葬的风俗，堡垒考古的清晰度——结合在一起，意味着我们能够相对较好地了解很多历史背景下的战争的不同侧面。我们在对战争能力进行量化分析时面临的问题，更多的是来自概念方面的挑战，而不是资料缺乏。

人们尝试着度量战争能力，这就像战争本身一样古老。几乎所有开战的决策都包含某种对社会之间相对军事实力的评估（即使侵略者往往高估了自己的力量，而防御者往往低估了本方的实力）。20世纪时，有一系列军事方面的行家里手和门外汉都曾试图设计出一些演算法，供将军们预测战争结果。

这些量化者中的第一位，某种程度上也是最具影响的一位，是博学多才的弗雷德里克·威廉·兰彻斯特（Frederick William Lanchester）。

他不仅是英国最重要的汽车工程师之一，还写过一本关于空战的先驱性的书，提出了一系列微分方程，来预测空战的结果。自那以后，兰彻斯特的方程被发展成为量化战争消耗的通用法则。

兰彻斯特方程因为一些不切实际的假设曾反复受到批评。到20世纪七八十年代时，美国陆军退役上校特雷弗·内维特·杜普伊（Trevor Nevitt Dupuy）研究出了一套更为复杂的“定量判断模式”，须使用不少于73种变量。但最近10年，已经有一种简单得多，也令人信服得多的替代方法被设计出来了。

所有这些方法都是为了对潜在的未来战争进行量化，并且都受过历史上实际战争的数据的检验。在不同历史时期的社会之间，或者地理上相距甚远、历史上从未交战的社会之间比较战争能力，困难要大得多，但是社会发展指数所必需的。专业军事人员经常用儿童游戏“石头剪刀布”来形容作战系统是怎样运行的：A系统（比如说使用步枪的步兵）会强于B系统（比如说使用军刀的骑兵），而B系统强于C系统（只有大炮的炮兵）；但是与此同时，C系统又强于A系统。因为军事能力总是依赖于环境的（也就是说，建立武装部队是为了在特定的地理和政治条件下，与某种具体类型的敌人作战。善于对付某种类型敌人的部队，也许不善于对付其他类型的敌人）。这种因时间和空间不同而结果变化极大的比较，必然比能量获取或社会组织那种广泛的比较要抽象得多。

战争能力的比较，归根结底一定是对社会可调动的破坏力的衡量。“破坏力”是指社会可投入战场的兵员数量。这种力量又会因武器的射程和火力、可调动兵力的数量和速度、防御能力和后勤保障能力而发生改变。而且，这些基本的要素——都理所当然地在许多时候和许多地方得到了很好的记录——一定要与一些没有得到很好记录，但同样重要的因素结合起来。这些因素包括士气、领导能力、指挥和控

制能力、清楚地理解战略的能力、行动力、战术原则、组织学习能力，以及经济、后勤、思想和政治等方面更广泛的参数。

技术问题令人望而生畏，但自从19世纪末期以来，战争游戏的研制者们（既有专业军事人员，也有业余爱好者）就一直在努力想办法，要将扑朔迷离的复杂现实简化为可供比较的数值。总的来说，军事史学家菲利普·萨宾（Philip Sabin）的说法可能是正确的，商业化战争游戏通常试图复制过多细节，但与社会发展指数非常相似的是，这些游戏的重大贡献是使得假设变得清楚明确了。

一些游戏发行多个版本，模仿不同时间和地点的战争，为跨越时间和地点思考战争能力提供了良好的起点。例如，电子游戏公司GMT的系列游戏“历史上的大战役”，包括公元前第二个千年西南亚的战车大战、公元前3世纪和公元前2世纪罗马共和国的战争、同时期印度的战争和13世纪蒙古人的战争等各种战役变体。至少在战术层面上，游戏允许启人深思的比较，不过，像所有这样的规则系统一样，这个游戏最大的价值也许仍是系统在似乎运转不灵时提出的问题。

原则上讲，社会发展指数所需要的跨历史比较应当与实际历史背景下的比较没有差异，但实际上，随时间推移而发生的巨大变化会使情况变得极为复杂。譬如所谓的军事革命，就经常是设计出新的作战系统，使旧系统根本无法与之匹敌。这方面最著名的事例就是英国皇家海军的“无畏号”战列舰。这种装备了大量武器和更厚钢板的战舰，是英国于1906年投入使用的，目的在于让以前所有军舰都被淘汰——只是因为海军战术的变化，才使得这种新式武器融入了旧式战舰仍然重要的系统里。

甚至对现代最致命的武器——核武器——来说，也是如此。核武器的破坏力远比非核武器大得多，但它们并非无可匹敌地更具破坏性。核武器的威力是以千吨或百万吨TNT（三硝基甲苯）当量来衡量的，这一事实便充分说明了这个道理。

拥有核武器的国家的破坏力使先前历史上的一切都相形见绌。在1942~1945年3年的轰炸中，美国空军第八航空队总共才向德国扔下了70万吨炸药；而在1961年万圣节那天，苏联试验的单一一颗炸弹，即所谓的“沙皇炸弹”，威力相当于5 000万~5 700万吨炸药。到1966年时，单独一颗苏联SS-9Ⅱ型导弹就可以携带相当于2 500万吨炸药的弹头，是美国“二战”期间轰炸德国的全部炸弹的破坏力30多倍；到20世纪70年代时，苏联总共部署了255枚这样的洲际弹道导弹。

然而，核武器的破坏力仍然可以和常规武器在同一个天平上予以衡量，正如放射性尘降物的毒害作用能以拉德来衡量，并与毒害作用较小的化学武器和生物武器对比一样。而且，像1906年后建造的无畏舰级军舰一样，核武器也被融入了更广泛的作战系统中，而这种作战系统依然依赖于1945年之前就在使用的武器类型（尽管以高得多的效率形式）。核战争是不可想象的，但并非不可衡量。

社会发展指数在衡量战争能力方面最大的困难，是将2000年的武装力量与之前阶段的武装力量的关系予以量化。1900~2000年，战争能力的飞跃实在太巨大，难以衡量；而1800~1900年，战争能力的变化也可称之为飞跃，尽管幅度不那么巨大，但也同样难以衡量。

一方面，这意味着如果我们在计分系统内将最高分250分打给2000年的西方，那么在估计1900年的战争能力时，在百分比上就会有相当大的误差幅度，更不用说1800年或任何更早的时期了。另一方面，由于现代破坏力与以往的差距实在太巨大，1800年前的分数会极低，也就意味着就社会发展指数的实际分数而言，误差也会极小。如我们将看到的，我为这些问题提供的答案，意味着1600年之前甚至没有任何作战系统能得到0.2分（也就是说，不到当代分数的千分之一）；而1500年之前，甚至很少有作战系统能得到0.1分。战争能力像城市规模/社会组织一样，是能量获取的功能之一，一旦能量获取水平达到了每人每天100 000千卡，只需在边缘处有一些小小的变化，就会刺激战争

能力向上激增。衡量战争能力对社会发展指数的主要贡献，是强调了工业化的20世纪和21世纪与所有先前的社会之间的巨大鸿沟。

西方是军事霸主吗

20世纪的变革

对于现代西方军事实力有很多评估，我主要依赖的是英国国际战略研究所每年出版的《军事平衡》（*Military Balance*），该书提供国家投入、军力、素质和后勤等方面的数据。

美国自“9·11”事件后又加大了军备投入，但即使在此之前，美国的军力就已经傲视所有对手了。公元2000年，美国可以代表西方获得满分250分。有很多国家在武装部队服役的军人人数上超过美国，俄罗斯的核武库也比美国大了几乎一倍，但美国在战争所有其他维度上都占有巨大优势，已经远远打破了平衡。美国军队的装备和供给比任何其他国家都好得多，训练和领导也强于大多数国家，机动性也很强。美国有11个航空母舰战斗群，完全主宰着全球海洋，美国空军在天上也发挥着同样作用。美国的核弹头及其运载工具也更加可靠，一般认为比其对手俄罗斯威力更大。

一旦我们从2000年回溯到1900年，量化战争能力的最大困难就出现了。1900年的西欧武装力量的数据记录很好，也很容易获得，但为1900年的西方计算一个相对于2000年的西方的分数，就非常困难了，因为军事系统之间的差距实在太大了。

2000年的军队比1900年要大，尽管并没有大许多。在某些方面，基本武器是相似的——1895年投入使用的英国的李–恩菲尔德步枪，准确射程为大约500米，枪口初速为每秒733米，而M16步枪（1963年装备于美国陆军，但其改进版在2000年仍是常规武器）的准确射程为

550~800米，枪口初速为每秒948米。然而，差异却超过相似性：M16步枪每分钟能发射700~950发子弹，而李–恩菲尔德步枪通常每分钟只能发射二三十发子弹（在测试条件下，其发射纪录为每分钟38发）。一把普通的M16步枪或者AK47突击步枪，比1900年的最好的重机射速都要快得多（马克沁重机枪每分钟只能发射450~600发子弹）。而最早的大致相当于M16或AK47步枪的德国MP18轻机枪，直到1918年才投入使用。

军事史学家通常将现代大炮的出现定为1897年投入使用的“法兰西75”。那是一种75毫米口径的膛线加农炮，配有长驻退机，意味着炮手们不需要在每次发射后都添加炮弹。这种炮的发射速度可以达到惊人的每分钟15发炮弹，射程达7.5千米。现代更复杂的大炮射速要慢得多。美国陆军最新式的榴弹炮——2005年投入使用的155毫米口径的M777——每分钟只能发射2~5发炮弹；然而，这种钛炮非常轻，可以装载在飞机上，射程达到24~30千米，而且，当使用“神剑导引炮弹”时，其圆径概率误差可达到24千米内只有5米（也就是说，50%的炮弹将落在距目标只有5米的范围内）。自20世纪80年代以来，制导武器的革命已经使每一门现代大炮的价值都超过了1900年时的数十门大炮，而运输机械化、通信手段和电子控制方面的先进同样令人惊叹。

在海上，1900年最强大的武器是新式钢铁装甲蒸汽动力的战列舰（“战列舰”一词是1892年最早开始使用的），通常排水量达到15 000~17 000吨，航速达每小时30千米，可装载4门12英寸口径的大炮，能发射400千克重的炮弹，射程将近23千米。1906年后，无畏号不仅配备了更重的装甲，还增加了6门12英寸口径的大炮，而且航速还提高到了21节[注]。1911年后，各国海军又纷纷用石油代替了煤作为燃料。所有这些改变都具有革命性的结果，然而即便如此，所有这些军舰与当代美国尼米兹级核动力航空母舰（排水量100 000吨，最大航速每小时56千米，可航行20年无需添加燃料，可载90架飞机，打击范围在700千

米以上）之间的差距，使得20世纪早期所有连续类型军舰之间的差异都相形见绌。

20世纪战争革命最惊人的部分无疑发生在空中。飞机最早用于军事是在1911年，当时意大利在与土耳其交战时使用了轰炸机和侦察飞行。这些早期飞机与2000年最精密的军用飞机（即1989年投入使用的B-2隐形轰炸机）之间的差距，足以令人窒息。B–2隐形轰炸机的航程达到11 000千米，巡航速度为大约每小时900千米，且它实际上是无法探测的，能够穿透任何防空系统，发射精确制导的炸弹或者10兆吨以上TNT当量的核武器。

我们可以很容易地比较各个时期的武器火力、军队机动的速度和范围，以及武装力量的无数其他维度。例如，人们普遍认为从1900~2000年，大炮的威力提高了20倍；从1918~2000年，反坦克火力提高了60倍。但是为20世纪全部变化打一个具体的分数，却要难得多。

表5.1 自公元前4000年以来的战争能力（以社会发展分数衡量）

年代	西方（分）	东方（分）	年代	西方（分）	东方（分）
公元前4000年	0	0	公元前/公元1年	0.12	0.08
公元前3000年	0.01	0	公元100年	0.12	0.08
公元前2500年	0.01	0	公元200年	0.11	0.07
公元前2250年	0.01	0	公元300年	0.10	0.07
公元前2000年	0.01	0	公元400年	0.09	0.07
公元前1750年	0.02	0	公元500年	0.07	0.08
公元前1500年	0.02	0.01	公元600年	0.04	0.09
公元前1400年	0.03	0.01	公元700年	0.04	0.11
公元前1300年	0.03	0.01	公元800年	0.04	0.07
公元前1200年	0.04	0.02	公元900年	0.05	0.07
公元前1100年	0.03	0.02	公元1000年	0.06	0.08

公元前1000年	0.03	0.03	公元1100年	0.07	0.09
公元前900年	0.04	0.03	公元1200年	0.08	0.09
公元前800年	0.05	0.02	公元1300年	0.09	0.11
公元前700年	0.07	0.02	公元1400年	0.11	0.12
公元前600年	0.07	0.03	公元1500年	0.13	0.10
公元前500年	0.08	0.04	公元1600年	0.18	0.12
公元前400年	0.09	0.05	公元1700年	0.35	0.15
公元前300年	0.09	0.06	公元1800年	0.50	0.12
公元前200年	0.10	0.07	公元1900年	5.00	1.00
公元前100年	0.11	0.08	公元2000年	250.00	12.50

我选择的2000年与1900年的西方战争能力的比率为50:1，这样产生的1900年西方战争能力的分数就只有5分（与2000年的250分相比）。很显然，这个分数跟瞎猜差不多。如果选择100:1的比率，则1900年的得分为2.5分，也许同样是个说得过去的猜测。不过假如比率是25:1，1900年的得分就会是10分，就不大可能了。

这一误差幅度比整体社会发展指数的误差（见第七章）要大得多，但2000年西方战争能力的分数与所有较早阶段的分数之间的巨大差距，意味着我们可以很容易地将2000年之前的所有分数减半或翻倍，也不会使指数出现任何可以辨别的差异。表5.1、图5.1和图5.2运用我估计的数字，展现了自公元前4000年以来东方和西方战争能力的分数。图5.3和图5.4则展现了如果我们将对2000年之前的所有估计减半之后得到的分数。

在纵轴上使用对数标尺使得差异更明显，所以图5.2在对数－线性轴上显示了我计算的分数，图5.4以同样方式显示了修订后的数字（即2000年前所有时期使用的都是减半后的分数）。修订后的数字当然使20世纪的破坏力比我的估计大了一倍，但是将2000年以前的分数减半的主要结果，并不是增加了现代/前现代的反差，而是使得公元前100~

公元200年东西方之间的差异变得极小，以致无法衡量（而我的估计则相反，显示出罗马帝国的战争能力稍强于汉帝国）。结论必然是，对2000年与1900年之间战争能力的比率的任何合理的估计——无论是50∶1，还是100∶1，还是只有25∶1——对于更大的社会发展指数都没有什么差别。

图5.1 公元前3000~公元2000年东方和西方的战争能力（线性－线性标尺）

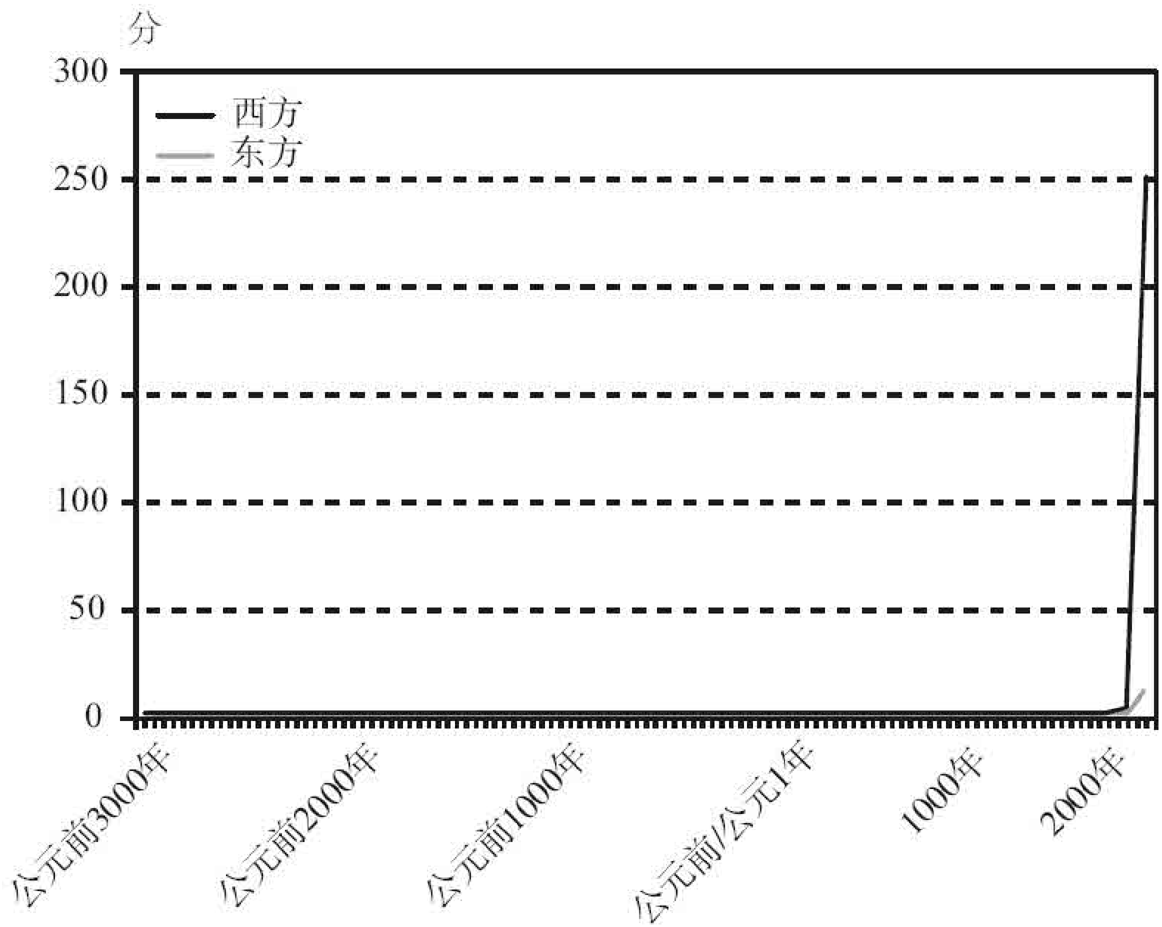


图5.2 公元前3000~公元2000年东方和西方的战争能力（对数－线性标尺）

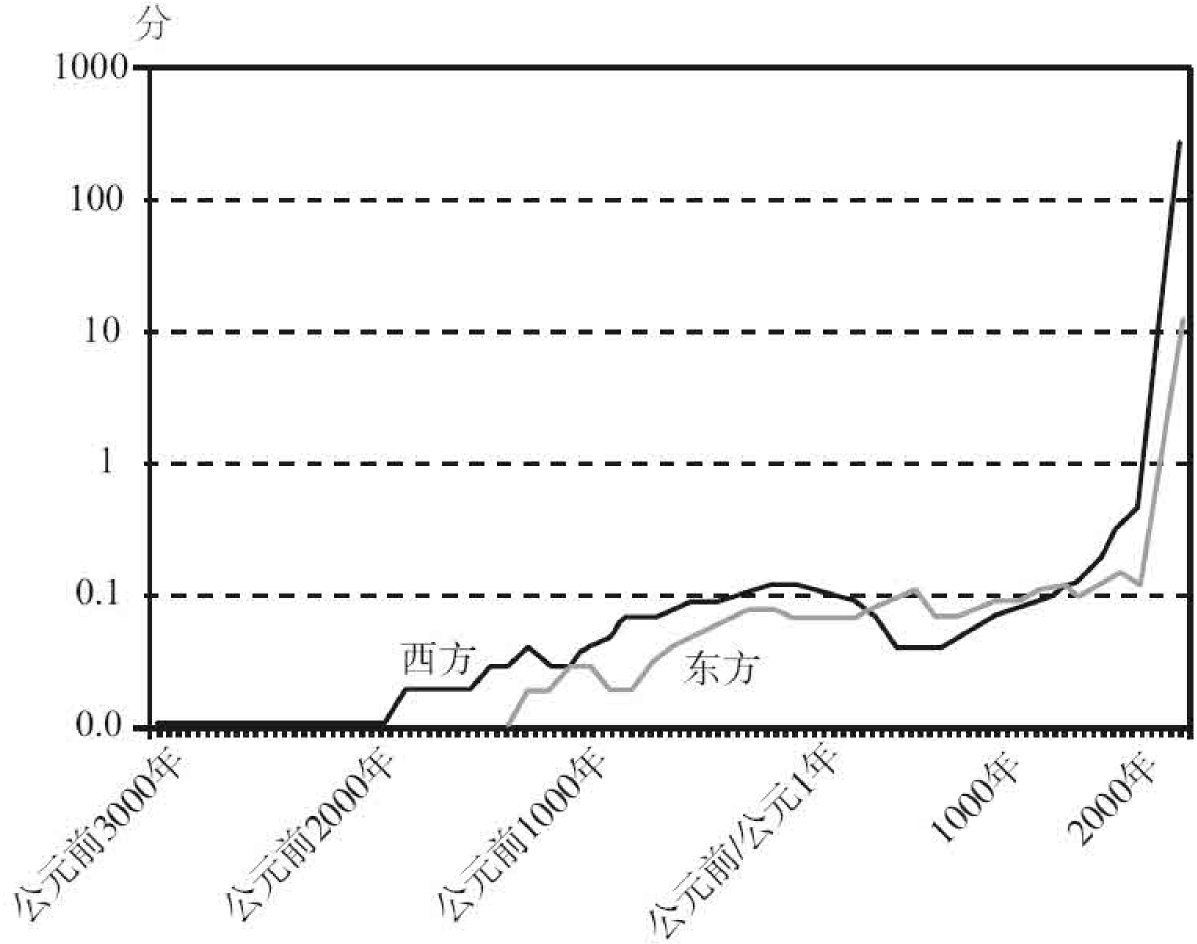


图5.3 公元前3000~公元2000年东方和西方的战争能力（以线性–线性标尺标绘，将2000年前的所有分数减半）

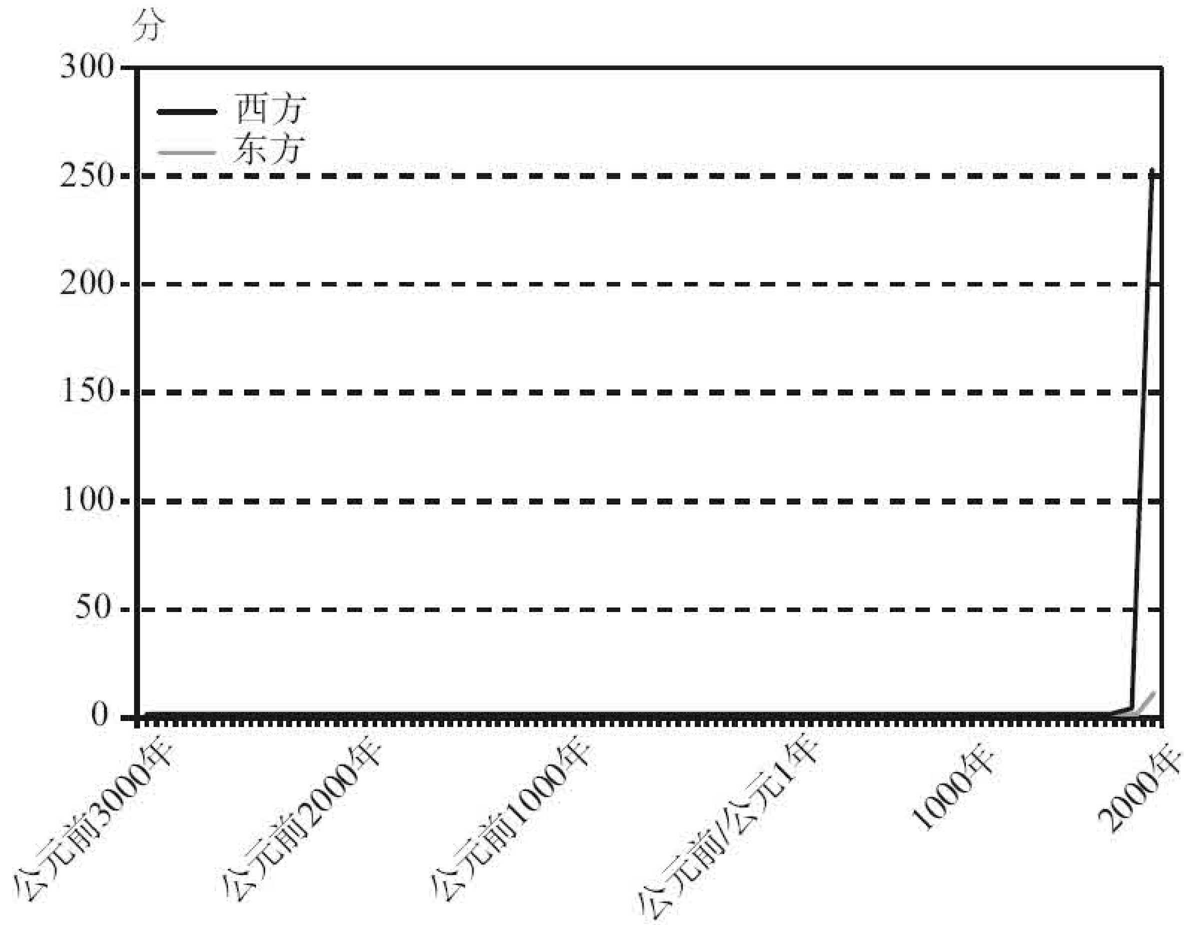
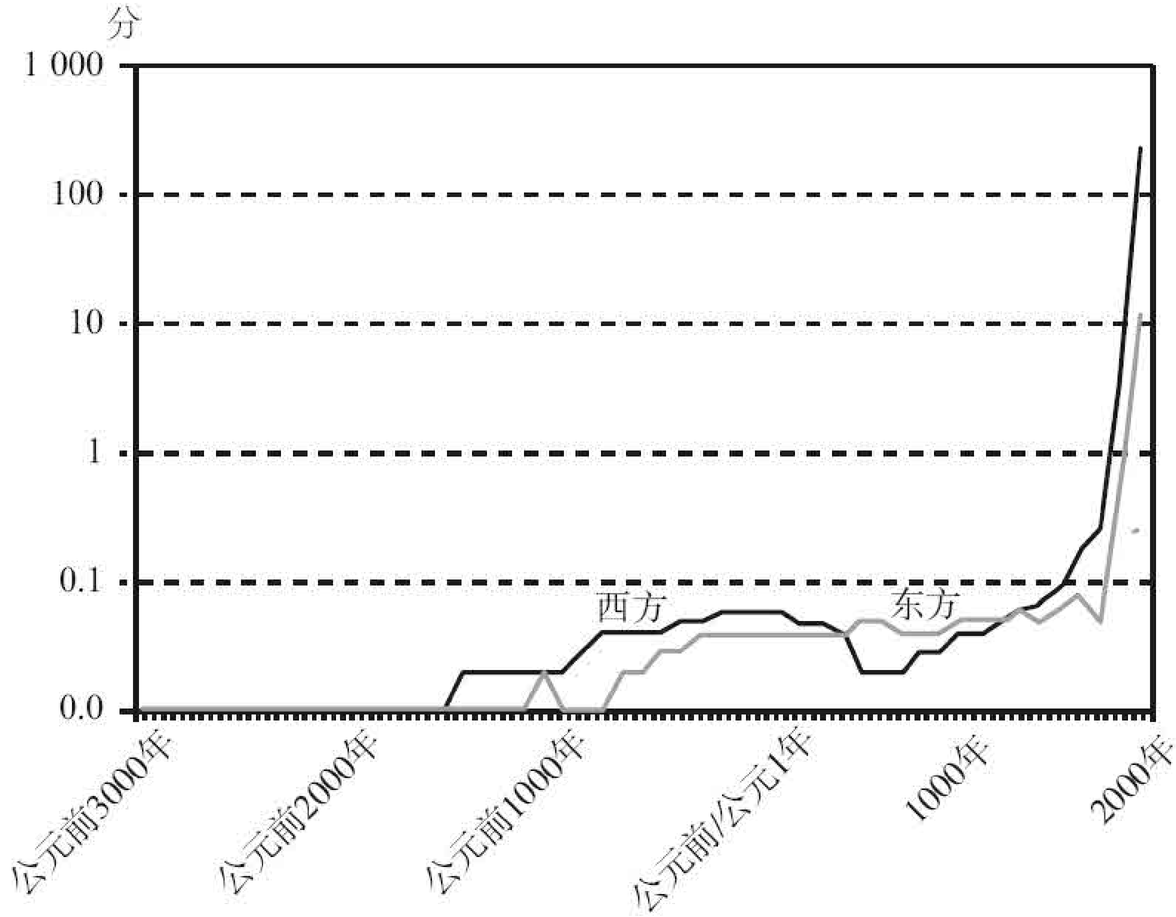


图5.4 公元前3000~公元2000年东方和西方的战争能力（以对数－线性尺度标绘，并将2000年前的所有分数减半）



1500~1800年欧洲的军事革命

1800~1900年西方战争能力的飞跃与1900~2000年的飞跃，根本不可同日而语，但也非常巨大。武器的射程、精度、射速，抛射的力度（因炸弹的发明而威力更大），军队的规模，运输的速度，以及后勤保障能力，在整个19世纪都普遍提高了一个数量级。然而，纯技术力量的大小取决于人们对其的应用。据军事分析家斯蒂芬·比德尔（Stephen Biddle）计算，1800年左右时，如果一个拿破仑式的千人步兵团向另一个步兵团发起冲锋，针对每一个攻方士兵，守方大致能打两发子弹；如果一个世纪后重复同样的步兵团攻防，那么就将有200多发子弹射向每一个攻方士兵。不过，令人惊奇的是，从1815~1918年，军队在战斗中的伤亡比例实际上呈下降趋势，因为战术家们采用了新的作战方式，使得部队暴露在敌人直接火力下的机会降到了最低。

法国人于18世纪90年代发明了“全民动员”，使得军队规模扩大到50万人——将近1900年世界最大军队的一半了——但是其主要武器滑膛步枪比1900年的步枪效率就差远了。即使是训练有素的拿破仑步兵每分钟也只能打大约4发子弹。滑膛枪的子弹可以射到400米开外，但射程一超过50米（最多75米）就很不准确了，只打一个单发实际上没有任何用处，甚至在打不超过75米的目标时，也只有多人齐射才有命中把握。在18世纪的一次演习中，步枪手们射击60米外一个宽达30米的目标，只有不到一半人命中。

滑膛大炮，尤其是能发射12磅(注)重炮弹的大炮，每分钟能发射4~6发炮弹，有效射程达到500米，于1800年开始成为战场主宰，但滑膛炮仍然远不及1900年的膛线炮更有效；平射弹道的炮弹是直到19世纪50年代才普遍使用的。

1800年最好的军舰，比如英国皇家海军的胜利号（1765年下水），在顺风的情况下航速可以达到每小时15~17千米，但在恶劣天气下就要慢得多。胜利号载有104门炮，总共能发射大约1吨实心弹丸，射程大约为2 000米。该舰与无畏号出现之前其他拥有钢铁装甲、蒸汽动力、炮弹和鱼雷的军舰之间的差异，也是非常显著的。

又一次，将军事系统的复杂性概括为一个单一的分数，是件非常主观的事情，但我认为1900年和1800年西方战争能力的比率大约为10:1，1800年的得分为0.5分。这个猜测的误差可能像我对1900年的猜测一样大（或者说与拿破仑时代的步枪单发的准度差不多），真实比率也很可能是20:1。如果相对于2000年的战争能力，我对1800年和1900年都高估了，那么我们也可以不给2000年打250分，给1900年打5分，给1800年打0.5分，我们完全可以在给2000年打250分的情况下，给1900年打2.5分，给1800年打0.13分，就像图5.5（线性－线性图）和图5.6（对数－线性图）所显示的。但是即使像现在这样把1900年前的

分数大大减少，对整体社会发展指数所产生的影响也仍然是微不足道的，因为所涉及的绝对数值实在太小了。

图5.5 公元前3000~公元2000年东方和西方的战争能力（将以线性–线性标尺标绘，1900年之前的分数降低之后）

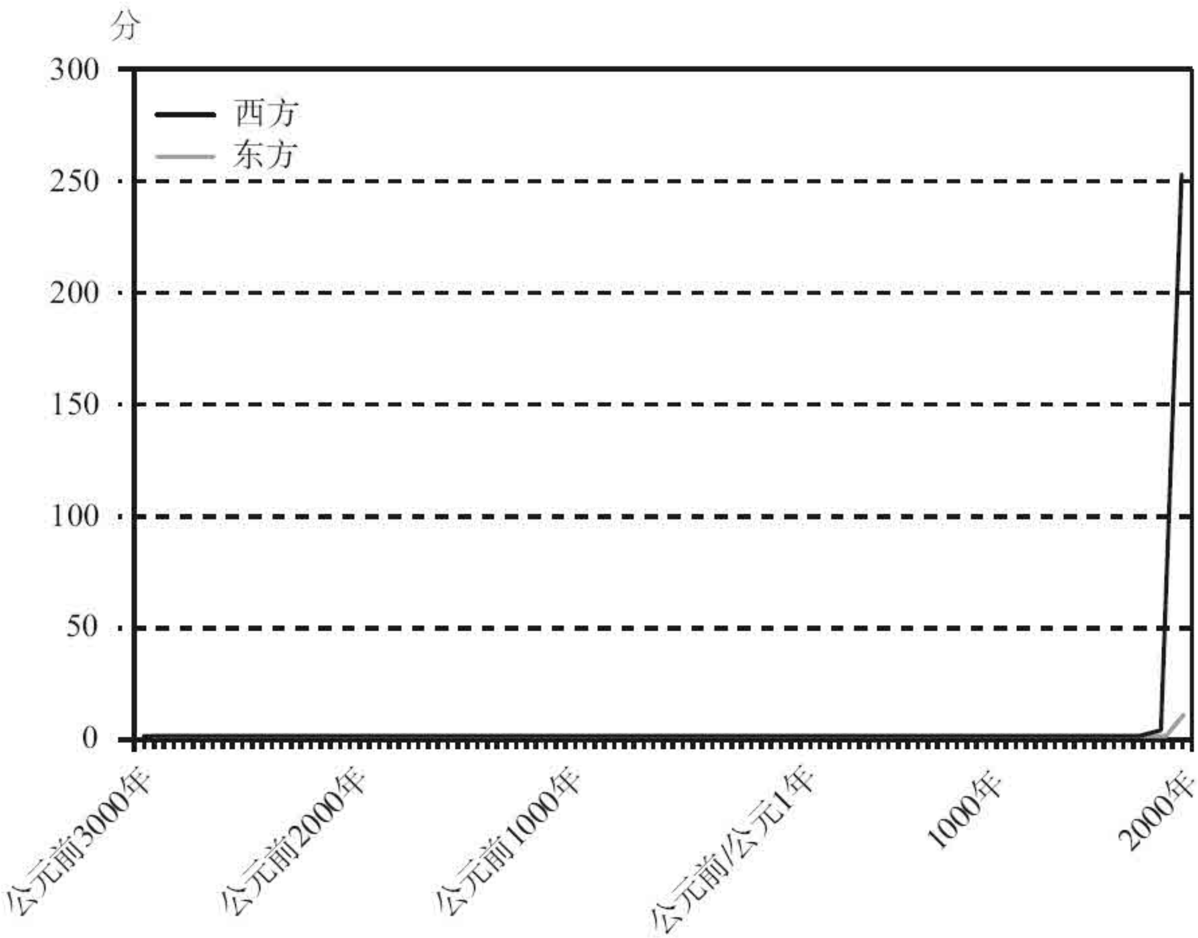
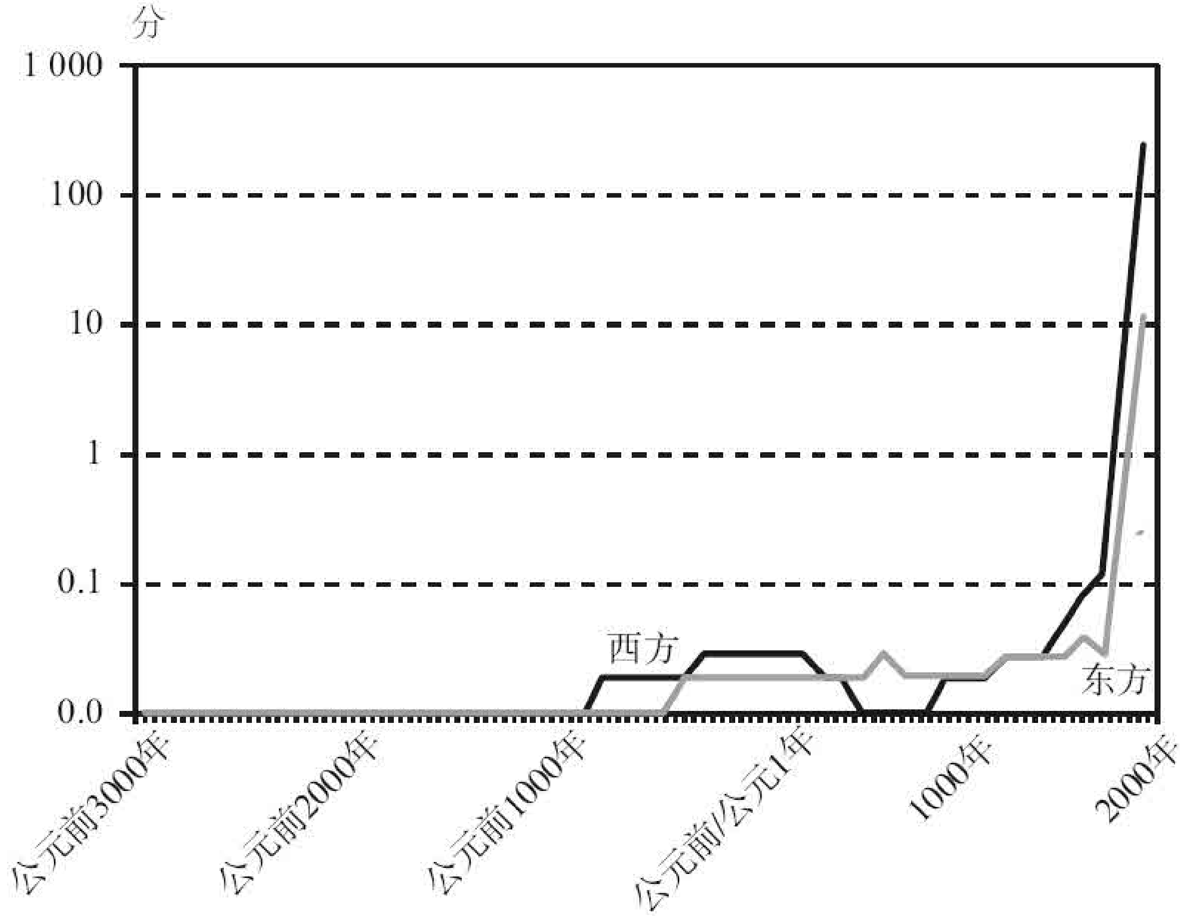


图5.6 公元前3000~公元2000年东方和西方的战争能力（以对数－线性尺度标绘，并将1900年之前的分数降低之后）



由于迈克尔·罗伯茨（Michael Roberts）以及最重要的杰弗里·帕克（Geoffrey Parker）的先锋性工作，1500~1800年这段时期已经广为人知是“欧洲军事革命时期”，其特点是陆海军规模、效率、火力和到达范围的巨大增长。然而与1800~1900年的变化相比，这些军事革命实际相形见绌，不过它们仍然将中世纪欧洲社会的军事能力远远地甩在了后面。

火器和组织形式发生了变化，而一些社会利用了这些变化，是促使军事革命发生的重要原因。使用火药的武器在14世纪20年代就到达了欧洲，但过了100年，它们才开始在陆地和海洋的战场上变得重要起来。甚至在1500年，火枪手们的发射率还在以几分钟打一发子弹，而不是一分钟打几发子弹计算，而他们的枪只在非常短的射程内才有效。特别是在英国，长弓在训练有素的弓箭手手里，每分钟能发10

箭，精度可达200米，不少士兵都不相信，弓将不再是最优越的武器。而在骑兵最为重要的大草原上，直到17世纪，弓仍然在战场上发挥着主导作用。

甚至早期的火绳枪也能发射比箭要重的抛弹（滑膛铅丸），因此有更强的穿透力，但火枪的主要优势是对枪手的技能要求较低，不像弓箭手那样需要学很多技巧。正如意大利的拉文那战役（1512年）、马里尼亚诺战役（1515年）和比科卡战役（1522年）所证明的，在适宜环境下，众多火枪是可以战胜弓箭和长矛的。早在1490年，威尼斯就决定用枪来取代十字弓，到了16世纪60年代时，英国人对长弓的喜爱已经确定无疑地被普遍认为不合时宜了。1594年，荷兰陆军发明了队列战术和火枪齐射，极大地提高了他们的效率（尽管他们的战术需要更多训练和管理）。17世纪30年代，古斯塔夫·阿道夫（Gustavus Adolphus）向世人展示了这种新办法的威力有多大。

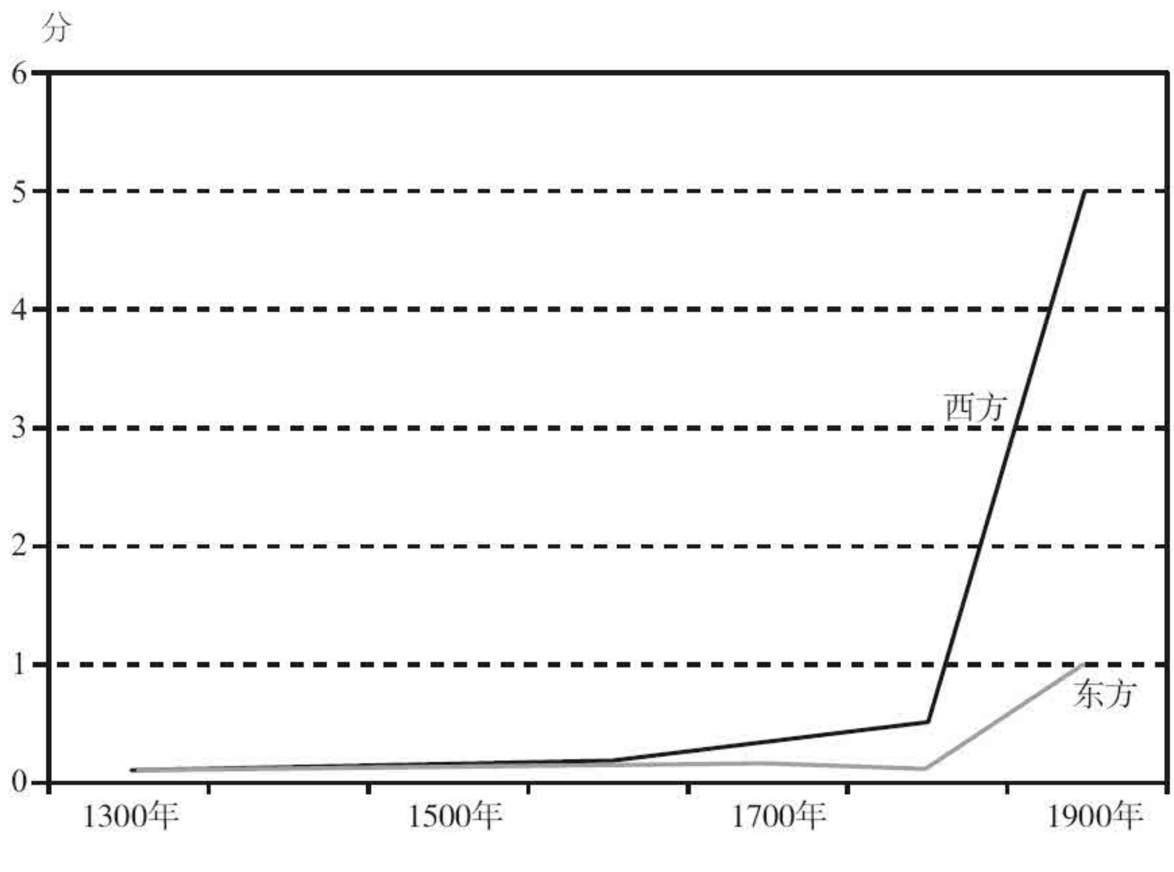
17世纪，燧石发火装置极大地提高了火枪的发射率。18世纪插座式刺刀的出现又使火枪手同时能发挥长矛手的作用。大炮的进步更是迅速。到1494年查理八世入侵意大利时，火炮已经使中世纪的石制堡垒失去了作用，但到17世纪中叶时，更复杂的土木工事又重塑了防守方的优势。

18世纪晚期组织方面的进步——尤其是法国在陆上发明了纵队进攻和军队的部门结构，英国在海上进行了战术革新——进一步提升了军队的绩效，但是最大的变革还是在组织方面。当时最强大的西欧国家法国，1500年时能够组织四五万人的军队；1600年时能组织8万；1700年时为40万；1812年拿破仑入侵俄国时率领的是60万大军。海军的增长要缓慢一些，1700~1800年，英国（当时最强的海上强国）、西班牙和俄国都将其现役军舰数大致涨了一倍，而法国的舰队在1689年路易十六入侵英国的计划失败后，实际上萎缩了。这段时期开始时，

奥斯曼土耳其的陆海军实力都还是西方最强的；但到结束时，军事力量的天平已经决定性地倾向了西欧方面。

将如此众多而复杂的信息转化为西方战争能力的单一分数，又一次需要动用非常主观的猜测，但是1500~1800年的变化，尽管有革命的性质，却明显比1800~1900年的变革要小得多（更不用说与1900~2000年相比了）。西方战争能力在16世纪大致增长了50%，在17世纪增长了100%，18世纪又增长了50%，在整个军事革命时期实力涨了4倍（与之对比的是，我估计19世纪涨了10倍，20世纪涨了20倍）。从1800年的分数0.5分向后推，1700年的分数大致为0.35分，1600年为0.18分，1500年为0.13分（见图5.7）。

图5.7 1300~1900年东方和西方的战争能力



从恺撒到苏莱曼（公元1~1500年）

大多数军事通史学家认为，在这段漫长的历史时期的前半期，西方的战争能力总体下降了，在后半期恢复。但意见并不完全一致，伯纳德·巴克拉克（Bernard Bachrach）在其一系列研究中，提出后罗马时代的西欧军队比其他历史学家设想的规模更大，更多地由步兵主宰，更善于远程作战，更多采用围困战术而不是直接战斗。然而，这只是极少数人的观点，我还是追随主流意见：西欧的军事力量自200年后开始下降，400年后下降得更快，600~800年时极其衰弱，然后缓慢恢复，1300年后复兴开始加速。倒是没有什么迹象表明14世纪黑死病肆虐后，在军事方面出现了堪与能量获取和社会组织相比的衰退。

然而，尽管战场上有重要变化，比如由于能够得到更大的马匹和马镫，使得重型骑兵兴起，伊斯兰军队中马上弓箭手的效率也在提高，但公元500~1300年战术的持续性更引人注目（实际上自公元前700年后的整个2 000年间，战术都没有什么大变化，因为普遍都在使用铁兵器和骑兵）。基本的军事要素——铁兵器、金属盔甲、步兵和骑兵结合的战术、箭术、围城机械、划桨结合风能的战船——在这段漫长的时期内几乎没有变化，真正的变化是在后勤和组织方面。

公元前1世纪30年代，罗马共和国大约有25万武装人员。他们被编制在效率极高的军团内，得到前现代世界最出色的后勤系统的支持。他们（大部分时间）由优秀的职业尉官和士官领导，即使他们的高级军官——特别是在共和国时期——有时不尽如人意。

公元3世纪的危机后，军队扩大了，在4世纪中叶时可能达到了大约50万人。关于晚期罗马军队的素质，存在很多争议，一些历史学家认为真正的问题是任务的性质变了，从边境防御转向了纵深防御，随之引起了组织上的变化，卫戍部队和野战军的差别越来越大，后者采用了比早期帝国军队更小的单位和更多的骑兵，所有部队也都更加依赖机动兵力。

有些旧说法称卫戍部队效率低下，可能言过其实了，但从2世纪60年代的安东尼瘟疫到378年的阿德里安堡战役，罗马军事力量恐怕的确严重衰退了（尽管还不是灾难性的）。

从阿德里安堡战役到609年波斯国王霍斯劳二世入侵拜占庭帝国，由于人口下降加上行政结构破碎，西方军队的规模和战斗力进一步下滑。到7世纪时，军队萎缩到只剩下几万人。阿拉伯人之所以得以迅速征服了波斯帝国和拜占庭帝国的大部，更多地要归因于帝国结构的瓦解，而不是哈里发的军事力量有多么强大。

整个西方中世纪，军队始终规模很小、组织涣散、供给不力，人数基本不到罗马帝国军队的1/10，效率更不值一提。中世纪的欧洲军队有学者仔细地研究，但没有得到那么透彻研究的拜占庭军队，特别是伊斯兰军队，在大约630~1500年也许仍很强大，尤其是在土耳其军队配备了成千上万名强悍的马上弓箭手后。

西欧的十字军于1099年奋力攻占了耶路撒冷，拜占庭军队也收复了一些失地，但总体而言，10~15世纪，优势在土耳其人一边。1527年，土耳其苏丹苏莱曼大帝据称能召集75 000名骑兵（大多数是弓箭手）和28 000名持枪步兵，另外还有野战炮兵。尽管他在1529年未能攻克维也纳，土耳其军队仍然是16世纪西方最强大的军队，很可能在17世纪还继续领先。同样，土耳其海军尽管在1571年著名的勒班陀战役中也吃了败仗，但直到1600年后的很长一段时间里，它始终是基督教势力在地中海上的一个强劲对手。

将所有这些历史简化为战争能力的分数，又一次需要将每支军队所面临的具体任务抽象化，但一些基本结论应该是合情合理的。1500年西方最大的军队的规模仍然比罗马共和国晚期或罗马帝国早期的军队要小得多，并且远不及罗马人技艺娴熟；但是火器威力的增长（尤其在攻打堡垒时，庞大的野战军与轻骑兵相结合，如奥斯曼土耳其军队），使我怀疑苏莱曼大帝的军事实力恐怕最终赶上了恺撒的军队。

如果给1500年西方的战争能力打0.13分，那么给1年打0.12分，是合情合理的。如果认为罗马军事力量直到4世纪时仍然很强，然后骤降的普遍看法是正确的，我们也许可以认为300年时有0.1分，到600年阿拉伯人征服前夕时暴跌至只有0.04分，到1200年时又反弹至0.08分，然后于1500年更快地攀升至0.13分（见图5.8）。[认为罗马的分数应该稍高一些（比如说0.13或0.14分）或稍低一些（比如0.10和0.14分似乎更为合理）的历史学家们，也可相应地调整300~1200年的分数。]

这些数字与历史文学中的定性估计是相一致的。然而，其中也包含了各种各样的抽象性和主观判断，反对者也许不会接受。这就是说，图5.9展现了也许最重要的一点：所有前现代时期战争能力的分数，包括恺撒时代和苏莱曼时代的分数，以公元2000年的视角看时都太小了，对它们进行任何可以想象的调整，都不会对社会发展指数造成多少影响。而且这还不是因为我们自己这个时代的军事力量超乎寻常而造成的人为假象。图5.10显示，即使从1900年的视点来观察，1~8世纪西方军事实力的变化都仍然小到可以不计。只有当我们从1800年的视角（如图5.11）回望时，我们才能看出早期分数间明显的高下。即使我们将600~800年的分数翻倍，或者断定罗马的战争能力直到1600年而不是1500年才被超越，那也没有什么不同。

图5.8 1~1500年东方和西方的战争能力

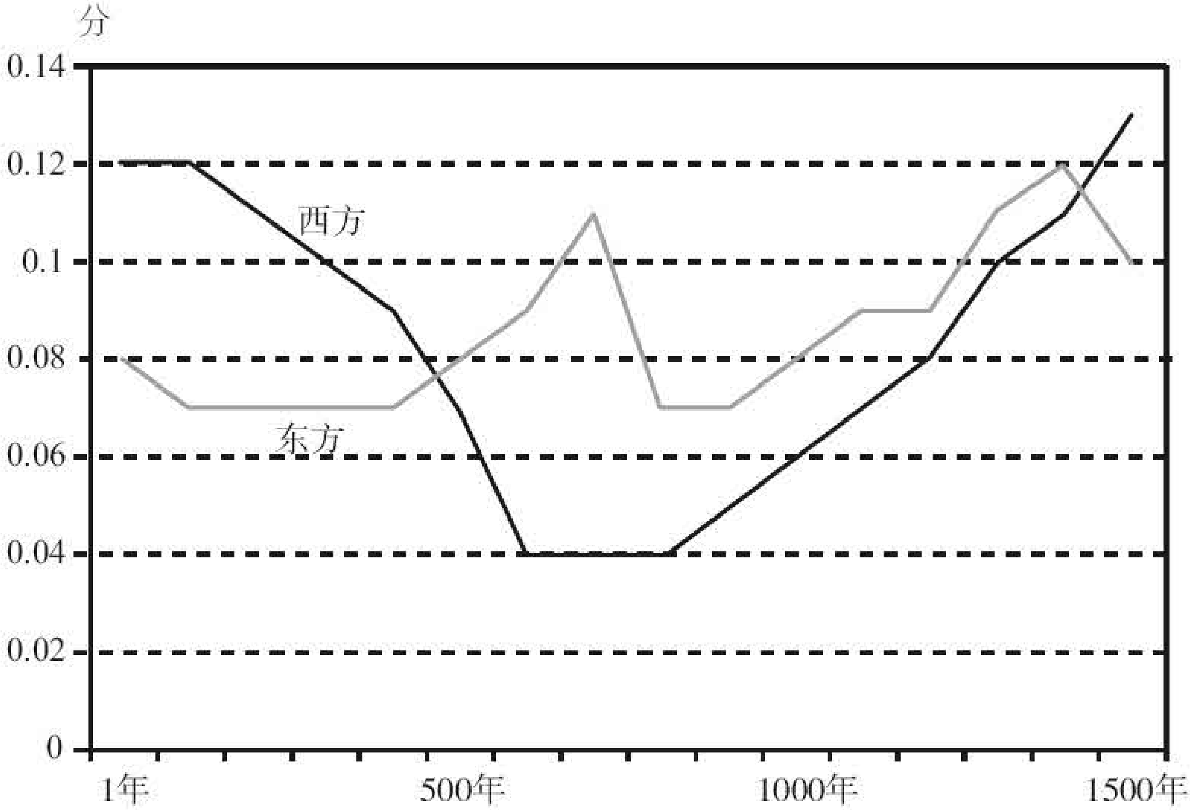


图5.9 1~2000年东方和西方的战争能力

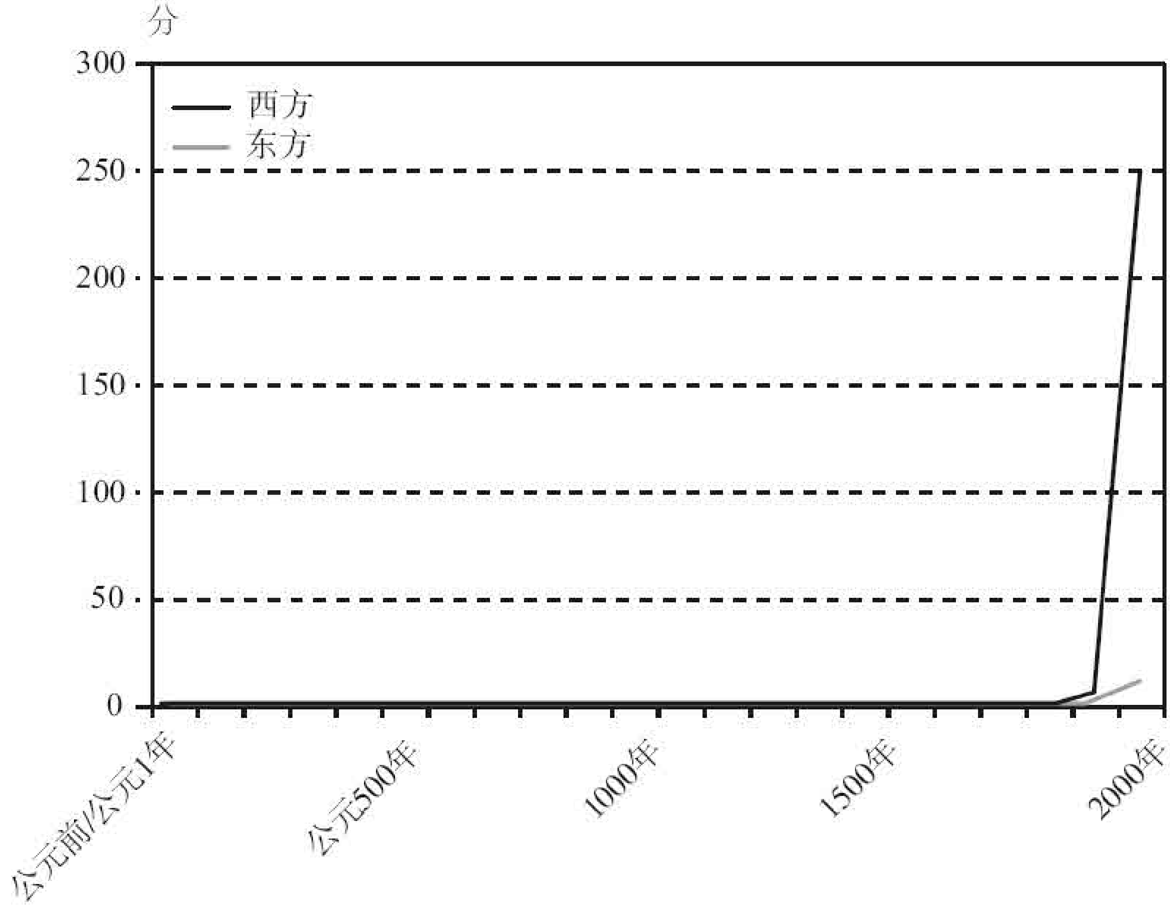


图5.10 1~1900年东方和西方的战争能力

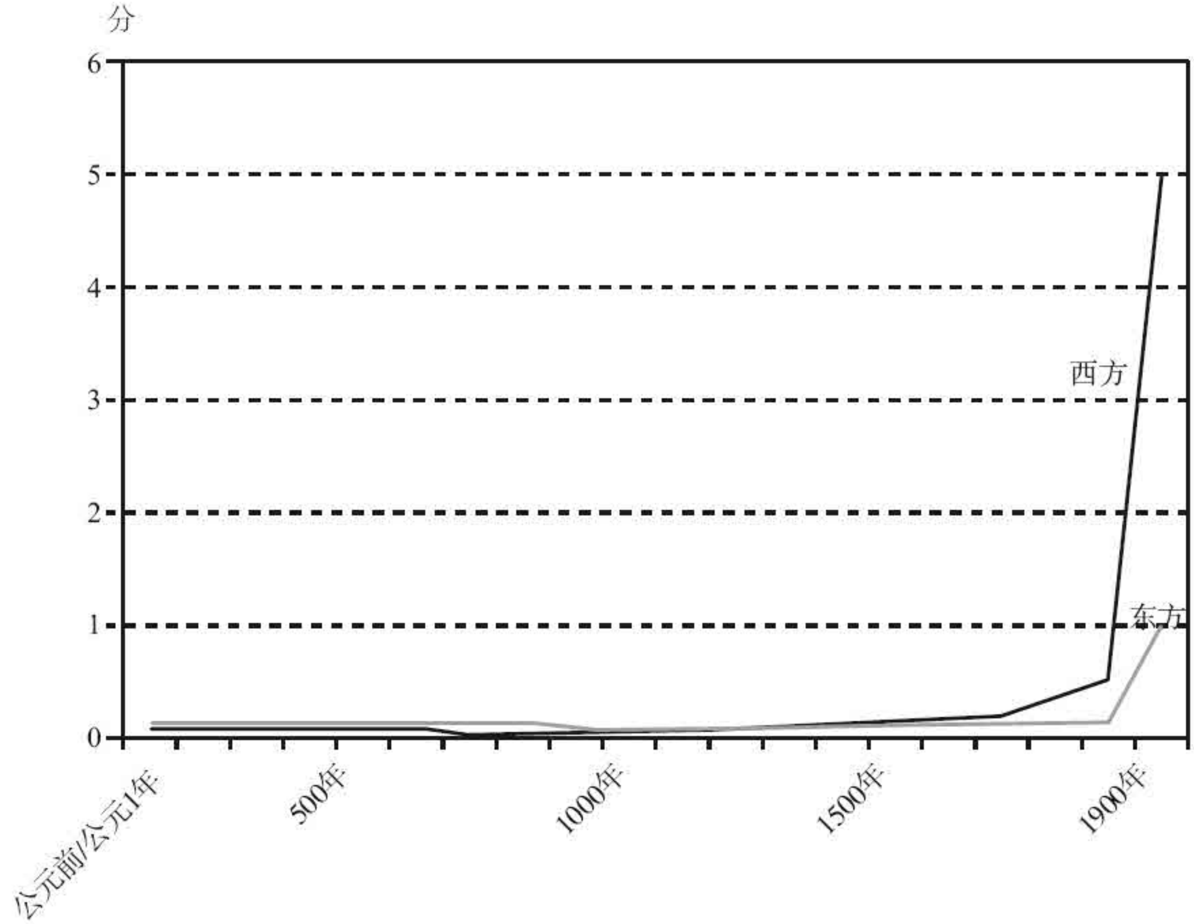
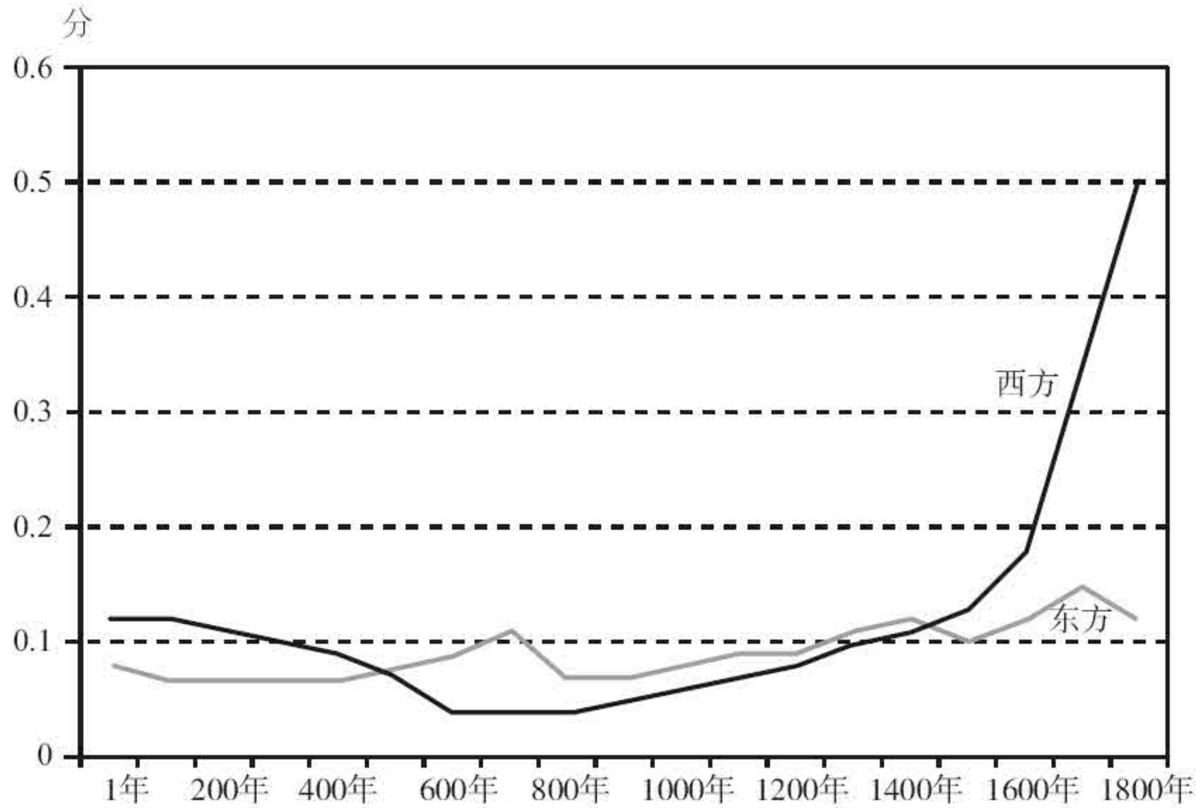


图5.11 1~1800年东方和西方的战争能力



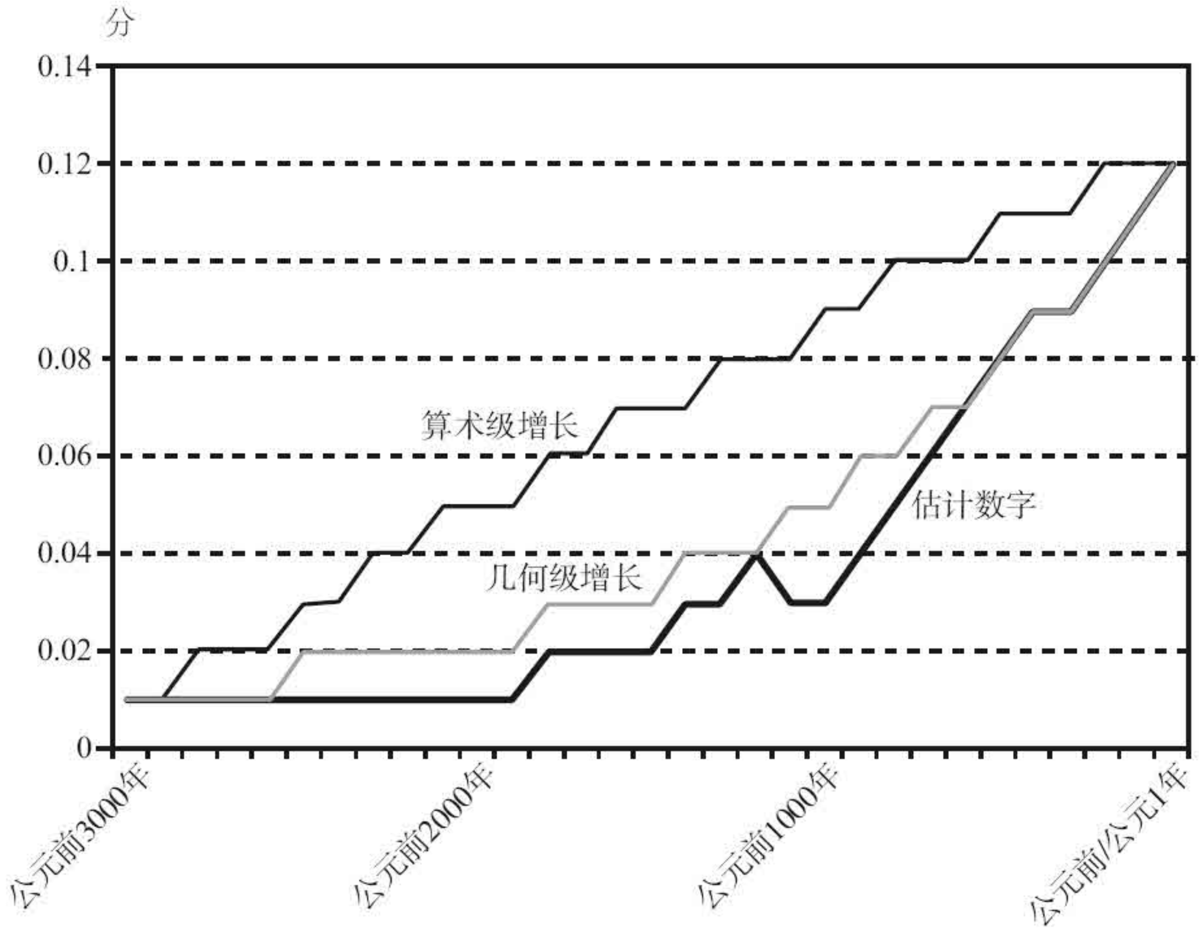
早年的战争（公元前3000~前1年）

公元前最后三个千年，将我们从埃及第一位法老纳美尔的时代，带到罗马第一位皇帝奥古斯都的时代，人类的战争能力有了相当显著的增强。在这段漫长的时期内，战场上的主要进步也许要列上：公元前第三个千年时，青铜武器取代了石制武器；公元前2500年，重型步兵兴起；公元前1600年左右，马拉战车普及；大约同一时代，合成弓（反射弓）取代了单弓（整片弓）；公元前1100年后，铁制武器取代了青铜武器；公元前900年左右，骑兵出现；公元前700年后，三桨座战船普及；公元前600年后，方阵战术兴起，并不断得到改善；公元前400年后，扭力弹弩和更大的战船（四桨座和五桨座的大帆船）出现；公元前300年左右，堡垒得到完善；公元前200年左右，更灵活的步兵战术得到发展。

对于军队组织方面的进步，我们也可以列出同样的单子：专门的战场阵形最早的迹象出现在大约公元前2500年，至今所知最早的常备军出现在大约公元前2350年，职业战车手出现在大约公元前1500年，由税收支持的常备军出现在公元前750年之后，专职水军出现在公元前500年之后，罗马人对征兵制的改革发生在公元前400年之后。军队规模也显示出同样的上升途径：大约公元前2350年时，阿卡德国王萨尔贡号称麾下有5 400人；公元前1274年加低斯之战时，交战双方分别投入了大约30 000名步兵、5 000辆战车；公元前845年，跟随亚述国王萨尔玛那萨尔三世出征的有10万人；公元前480年，波斯入侵希腊时出动的也是10万大军（准确的数字有争议）；公元前3世纪40~60年代罗马和迦太基大战时，动用的水军兵力也以10万计。

图5.12展示了3种用数字表示公元前3000~前1年战争能力的结果。我给公元前3000年之前的战争能力打了0分，并不表示我支持曾经一度流行的一种观点：国家出现之前的社会都是和平之地。这种理论已经被彻底否定了。0分纯粹是技术上的问题，反映的是社会可用的破坏力太小，发动不了足以记录在社会发展指数上的战争。

图5.12 估计公元前3000~前1年西方战争能力的3种结果



我们当然可以从其他假设起步，比如说，从我们至今听说的第一支常备军出现在美索不达米亚的时间，公元前24世纪开始计分，但没有明显理由重视这些假设中的某一个而忽视其他。我从公元前3000年开始计0.01分，纯粹是因为这是一个便利的整数，但是也没有其他看似合理的假设能使社会发展指数显示出任何比较明显的差异来。

图5.12中最上方的线，显示了公元前3000~前1年按简单的算术增量增长的战争能力，中间的线显示了按几何量级增长的战争能力（即以每世纪8.65%的稳定增长率增长），最下方的线显示的是我对变化率的估计。（从公元前3000年的0.01分到公元前1年的0.12分，算术曲线和几何曲线都不是平稳上升的，因为涉及的数字实在太小，而最低分数是0.01分，上升线都不可避免地是曲折的。）

算术增长显然不符合实际。那意味着到公元前2200年时，阿卡德国王沙尔卡里沙里和埃及法老佩皮二世的军队（大部分人有青铜武器，但也有一些士兵只有石制武器，步兵几乎没有盔甲，军中没有战车或骑兵，只会构筑非常初级的防御工事）可以得0.04分，简直与伍麦叶王朝和阿拔斯王朝的哈里发的军队（有铁制兵器、反射弓、骑兵和骆驼部队，会构筑坚固的城堡）一样强大了。那还意味着公元前1300年时埃及法老拉美西斯二世的军队（可得0.08分），像公元6世纪东罗马帝国皇帝查士丁尼的军队一样威武了。这两种结论都是绝对不可能的。

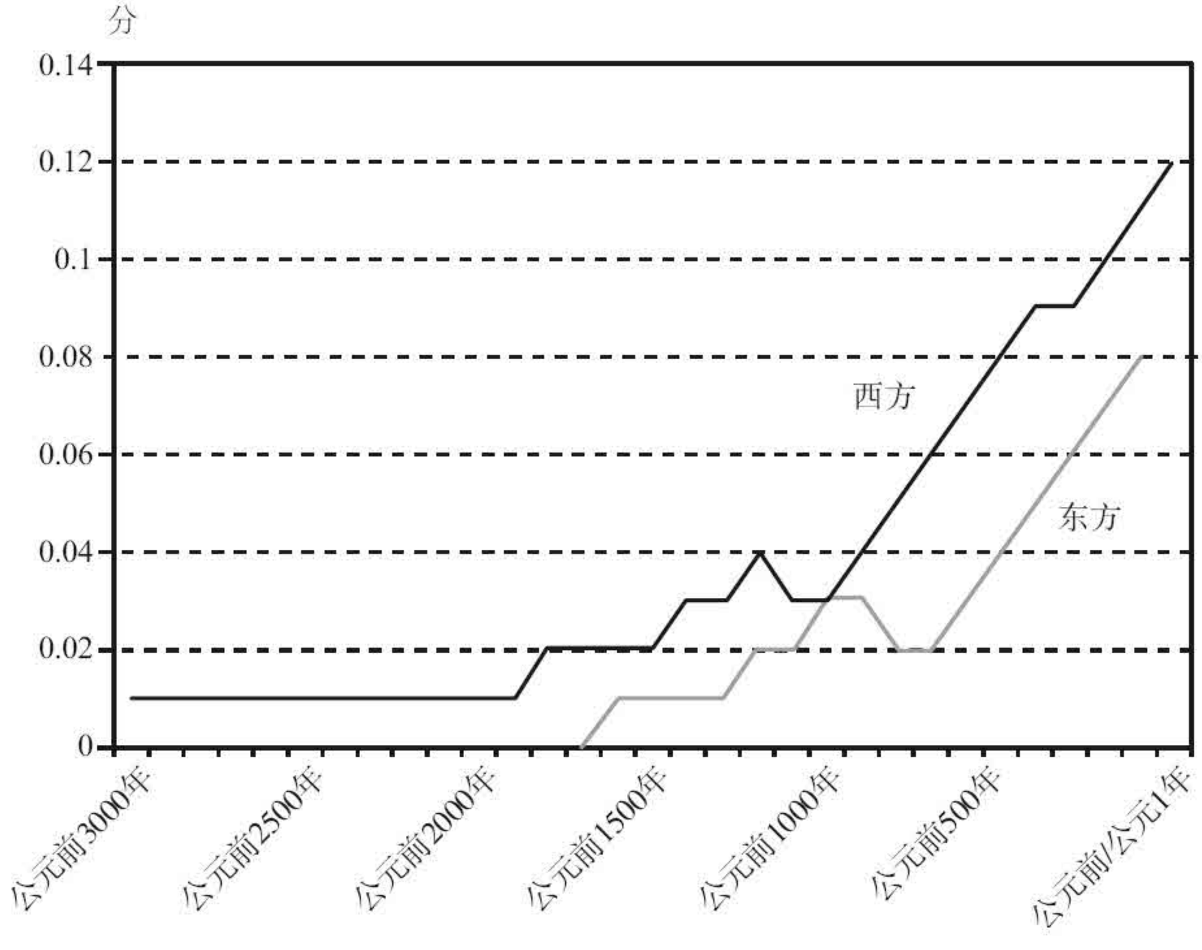
几何曲线似乎要可信得多，尽管该曲线掩盖了公元前1200~前1000年的崩溃，肯定过于简化了事实。公元前2200~前2000年的崩溃也对战争能力有巨大影响，但是又一次因为那个时期的分值太低（只有0.01分），无法在图中将衰退显示出来，除非我们假设公元前2100年美索不达米亚和埃及的战争能力跌回到国家产生之前、青铜利用之前的水平，这似乎也是不可能的。

我估计的增长率走向与简单化的几何曲线相差甚远。我的曲线在公元前3000年后缓慢起步，在公元前1200~前1000年的黑暗时代有所下降（从0.04分降至0.03分），继而在公元前第一个千年早期有一次快速增长。（无论在几何曲线上还是估计曲线上，公元前400~前300年的分数都是相同的0.07分，倒不是因为这段时期战争没有发展——这个世纪的军事水平从伯罗奔尼撒战争时期的重装步兵和三桨座战船发展到了亚历山大大帝和迦太基时期的各种武器协同的战术和五桨座战船——而是因为要把极小的数字凑成整数：公元前400年的分数恰好勉强超过0.07分，而公元前300年的分数不足以进到0.08分。）

几何曲线和估计曲线都显示，公元前13世纪，也就是国际时代的国王们似乎正要将东地中海地区变成一个单一大帝国的时期，当时的战争能力大致处于公元7世纪所跌回的相同水平（0.04分）。公元7世

纪时，拜占庭帝国和萨珊波斯都在分崩离析，阿拉伯的征服者们正企图夺回他们先前的领地。估计曲线还显示，古代的战争能力在约公元前900年时恢复到公元前13世纪的水平，当时诸如阿达德尼拉里二世等亚述国王也建立起大帝国。在《西方将主宰多久》一书中，我提出了这些看似合理的结论。最后，我的估计还表明，公元前200~公元200年罗马的战争能力与公元1300~1500年西方的水平非常接近——中世纪晚期的欧洲人应该会认同这个观点。

图5.13 公元前3000~前1年东方和西方的战争能力



难以评估的东方

2000年东西方的军事平衡

2000年东方最大的军事强国是中华人民共和国，不过尽管获得其军事力量的大体数字非常容易，给东方的战争能力打多少分却困难得多。

2000年时，美国与中国在市场交易率方面的投入比超过20:1，在购买力平价率方面超过9:1，美国核弹头数量对中国的优势超过25:1，洲际弹道导弹超过10:1，核潜艇为14:1，航空母舰编队为11:0。两军主战坦克的数量大致相当，但美国坦克的质量远比中国好得多。在从卡车到直升机在内的所有其他武器上，美国都占压倒优势。就总体技术能力而言，美国领先得甚至还要多。西方的军事主宰当然不是全面的，经常有分析家对美国海军是否敢在台湾海峡与中国众多潜艇和反舰导弹直接对抗表示怀疑；但中国在其近海之外几乎没有施展其军事力量，而美国却像个巨人一样跨坐在其他国家上。

公元2000年西方的战争能力显然比东方强很多，但究竟强多少？据我所知，很少有人试图把这个问题简化成一个单一分数。最广为人知的数字比较恐怕要数国家实力综合指数（Composite Index of National Capability，简称CINC）了。这是国际关系方面广为应用的一个计分系统，旨在标示每个国家的硬实力在世界上所占的百分比。其分数可回溯至1816年，但只能用于进行国家之间的共时比较，不能衡量历时性的能力变化。

不过，CINC所衡量的国家实力比我此处要查究的战争能力要广泛得多。该指标给每个国家打分，是根据其人口规模、城市化程度、钢铁产量、能量消费水平、军费开支，以及其军事人员的总数。按照CINC，2000年时，中国（16%）已经超越了美国（14%），尽管两个大国之间还存在着巨大的军事不平衡。

战争游戏设计者詹姆斯·唐尼根（James Dunnigan）在他的《如何作战》（*How to Make War*）一书中，采用了非常不同的办法，给不同国家的“战斗力”打不同的分数。他对陆上实力和海上实力分别打分，

在两个项目上都将美国定为榜首。在陆上，美国得2 488分；中国排第二，得827分。在海上，美国得302分，中国排第5，得16分。（英国排第二，得46分；俄罗斯排第三，得45分；日本排第4，得26分。）

如果我还继续沿用在社会发展指数中使用的技巧，只关心东方和西方最发达的地区，那么采用唐尼根的数字，将使公元2000年时西方和东方战争能力的比率大致为：陆上3∶1，海上19∶1。如果我们将陆上和海上的分数相加，那么美国就是2 790分，中国为843分（比率为3.3∶1）。如果我们改变一下，对陆上和海上实力等量齐观，将美国在两个领域的分数都转换成125分，同样合计入我在此使用的250分系统内，那么中国的得分为48.17分（陆上41.55分，海上6.62分），西方和东方的比率则略高于5∶1。

唐尼根没有解释他的分数是怎么得来的，但如果西方和东方战争能力的比率在3∶1~5∶1这个范围的话，必然要设想投资收益是递减的，因为美国的军费开支远远超过中国，2000年时双方的比率为9∶1~21∶1。在技术的复杂精密性方面也应大大加权，因为美国占有巨大优势的是复杂的技术密集型武器，如洲际弹道导弹、反导弹系统、隐形轰炸机、精确制导导弹和航空母舰，而不是诸如冲锋枪和手榴弹之类的简单武器。美国在电子战方面领先多少，还须继续观察，但2010年和2012年其分别利用“Stuxnet蠕虫病毒”和恶意程序“火焰”发动的网络攻击分别获得成功，表明美国的优势相当大。

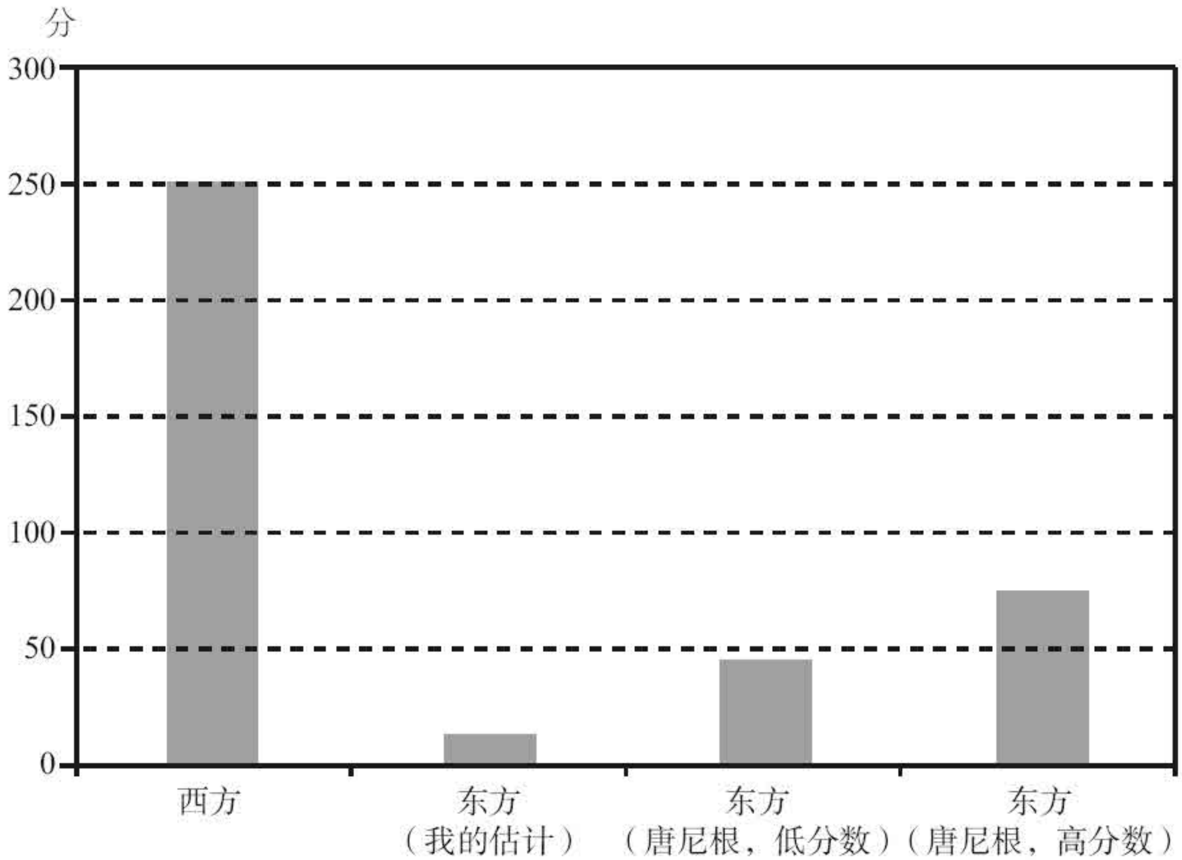
美国及其盟友在击败伊拉克和阿富汗的低技术敌人时所遇到的困难，表明唐尼根的假设有其价值，但也有证据显示，这些困难既源于战略和原则的失误，也与西方战争能力固有的局限性有关。还有一些军事分析家认为，投资实际上是收益递增的，在信息处理技术不断完善和信息传输系统越发精确的驱动下，军事革命已经给作战方式带来了极大改变，程度不亚于现代早期欧洲的军事革命（速度却要快得多）。1991年和2003年对伊拉克常规部队“一边倒”的战争，证明这种

观点也有其价值。军事革命似乎极大地改变了传统国与国之间交战的方式，扩大了西方相对于世界上其他国家在战争能力方面的领先优势，但对于占领和安抚战败国却没有多少影响。

公元2000年的西方和东方战争能力的对比大致为20∶1，比唐尼根提出的3∶1~5∶1要高得多。这将是历史上西方和东方战争能力最为悬殊的比率，甚至使19世纪和20世纪的比率都相形见绌，但是，公元2000年时东方和西方军事力量方面巨大的技术差距似乎证实了这一点。

如果我的估计是合理而准确的，在社会发展指数上，与西方的250分相比，2000年东方的战争能力只能得12.5 分。如果唐尼根的估计更准确，那么2000年东方可得48.17分（图5.14中估计的“低分数”）或75.54分（“高分数”）。

图5.14 对公元2000年东方和西方军事力量对比的不同定量估计



东方的现代军事革命（1850~2000年）

19世纪中叶，现代西方作战系统传到太平洋地区，是东方军事史上意义最深远的裂变。中国军队使用火器比西方时间更长，但没有跟上西方自15世纪使用火药武器后的前进步伐。日本则相反，使用火器相对较晚，却在16世纪成为变革的中心。然而，日本统一并放弃了丰臣秀吉的扩张政策，意味着日本军队自17世纪早期到19世纪早期没有打过大仗，于是日本火器发展在这段漫长的时期停滞了。

1840年和1850年，当西方的海军远征舰队分别来到中国和日本海岸后，中日两国均仿效起西方的军事实践来，但日本对新挑战的适应要成功得多。日本政府于1873年引进了欧洲式的征兵制度，并于19世纪70年代末废除了武士制度，继而先后按照法国和德国的路线建立了陆军，按照英国的路线建立了海军。1880年时，日本仍然远远落后于西方列强，陆军只有71 000人（只相当于德国陆军人数的1/6），海军舰船总吨数只有15 000吨（相当于英国海军的1/40）。但到1900年时，日本陆军有234 000人（几乎相当于德国陆军人数的1/2），海军舰船总吨数达到187 000吨（几乎相当于英国海军的1/5）。

日本军队的素质也得到了极大的改善。1894~1895年他们在侵略中国时，表现出了对西方军事思想、纪律、组织（和硬件）的精通。1900年，日本军队在围剿北京的使馆区的义和团时扮演了主要角色。1902年，英国认识到与日本结为海上同盟，是保证其在太平洋事务上继续发声的最好办法。1904~1905年，日本在对俄国的战争中取得了令人震惊的胜利（尽管这场战争差点导致日本破产）。虽然日本的战争能力比欧洲所有主要强国要弱很多，但已经成为一个地区性强国了，也许是当时世界上唯一能抵御欧洲人武力的非西方国家。

日本在1914~1915年和1941~1942年战争中的惊人胜利，都是在西方列强耽于欧洲战争而无暇旁顾的情况下取得的。日本在《凡尔赛条约》中得到了它觊觎的大部分利益（尽管它在条约中加入主张种族平

等条款的要求落了空），然而当它在1942~1945年不得不面对美国的认真抵抗时（即使美国只是将太平洋战场列为二级前线），东西方战争能力的持续差距就彰明较著了。

日本在1945年被几近全面地解除了武装（尽管到20世纪末时其海军又一次成为一支不可小觑的地区性力量），但中国在1949年结束其内战后，重新成为东亚强国。中国在1950年的朝鲜战争中发挥了巨大作用（虽然也付出了惊人代价），在1962年的中印边界自卫反击战中取得胜利，在1964年成功试验了其首颗原子弹。然而，人民解放军的训练和专业化在20世纪60年代的“文革”中遭受了严重损失。在20世纪70年代，其组织、理论和装备方面的严重缺陷暴露无遗。邓小平于1979年启动了军队现代化改革。20世纪90年代，中国军费开支开始大幅增长，在10年间涨了4倍，在接下去的10年内又涨了4倍，尤其强调加强海军，并对美国及盟友潜在的压倒优势形成不对称反应。到21世纪20年代时，中国的军费开支也许会赶上美国，但在2012年时，东西方的军事差距仍然是巨大的。

东方的战争能力在整个20世纪一直落后于西方。日本军队在20世纪40年代初期与西方军队的战斗中取得了显赫的战绩。中朝军队和越南军队在20世纪50年代和60年代与欧洲人和美国人的战争中也占据了上风。但在所有这些战争中，东方军队都得益于一个现实，就是在西方人的眼里，这些战斗都发生在争夺欧洲的大战争中的二级战场，西方的重点先是与纳粹德国的争夺，继而是与苏联的争夺。东西方战争能力的差距在1900~1940年有所缩小，但仍然很大，而在接下去的60年中又进一步拉大了。

为1900年东西方军事力量的对比打一个分值，没有2000年那么难。如前面提到的，1900年时德国陆军对日本陆军的优势在2:1以上，英国海军对日本海军的优势将近6:1，而且无论陆海军，欧洲军人的素质也要高得多。我估计1900年时西方和东方之比大致为5:1。1900年西

方战争能力的分数已被定为5分，意味着东方的得分是1分。还有更深的含义，东方的战争能力在20世纪增长了12.5倍，而西方增长了50倍，东方2000年的军事力量比西方1900年时高出2.5倍。然而，假如我们采用唐尼根估计的数字，也就是东方的军事能力在2000年为48.17~75.54分，那么我们就不得不相应接受一个更高的增长数字（即东方的战争能力在20世纪增长了50~75倍），这似乎太过分了。

火药时代的东方战争能力（公元1500~1850年）

新近的一系列研究使得中国2 000年帝制以来历史上关于东方战争能力的大致轮廓已经相当清楚了。主要的难题又一次是打一个精确的分数，但是在东方，涉及的数字（因而也就是合理的误差）在大部分时期甚至比西方还小。

将1900年之前东方和西方的战争能力直接进行比较，是件虽不精细但却仍然可行的事情。到1800年时，西方已经显然要强大得多了，也许在1500年时，也就是西方军事革命开始时，就已经更为强大了。明朝在必要的时候，的确有能力召集庞大的兵力（尤其是在15世纪上半期应对草原战争时），但没有像欧洲人一样有效地利用火药技术。

16世纪时，西方的大炮已经确定比东方的优越了。明朝政府也许早在16世纪20年代就见识过一些西方的大炮了，但即使这样，他们到16世纪40年代仍然对这种大炮很感惊奇。那时日本的军械师已经非常高效地仿制出一些这样的炮，不过数量仍然极少。甚至在16世纪中叶扭转了抗倭战争局面的著名的戚家军，拥有的火枪手与同时代的西欧军队相比也少之又少。他们的枪通常都造得很粗糙，往往会自爆，这使得枪手们不敢离自己的武器太近，因而也就不能有效地瞄准。戚家军的人数从来没超过10 000人，对海战的影响远比对庞大的明朝陆军的影响大。戚继光实施的一些新的海战备战措施是非常必要的。明朝水师自15世纪早期以来衰退得非常严重，手忙脚乱地付出了许多仓促的努力，才建立起一支部队，于16世纪90年代与朝鲜水师联手打退了

日本的进攻。陆军也同样如此。例如，北京的禁卫军到1564年才用铅制炮弹替代了泥炮弹，到1568年才（像欧洲人一样）改用了铁制炮弹；直到15世纪70年代，戚继光才采用了将轻型炮装置在由柳条栅栏围护的手推车上的办法，就像1444年匈牙利人在瓦尔纳与奥斯卡人作战时使用的那种车。

16世纪明朝的战争能力无疑比哈布斯堡帝国要弱得多（更不用说奥斯曼帝国了），某种程度上甚至比小小的荷兰共和国还要弱。我提议1600年给东方打0.12分（西方为0.18分），在丰臣秀吉发动朝鲜战争时，日本的军事力量已经与中国持平，东方1500年只能得0.1分（西方为0.13分）。这意味着中国的战争能力直到1600年左右才达到罗马巅峰时期的水平，尽管那时候人类使用火器已经有三四个世纪了。

中国的战争能力在整个17世纪是处于上升状态的。1696年，康熙与准噶尔游牧部落作战时，带了235门重型炮（每门重4~5吨）和104门轻型炮（每门重40~400千克）。但欧洲人的战争能力增长得更快。我估计1600~1700年欧洲人的战争能力翻了一番，从0.18分增至0.35分；这段时期东方的战争能力只增长了25%，从0.12分增至0.15分——意味着康熙的军事实力在罗马皇帝奥古斯都（0.12分）和哈布斯堡皇帝腓力二世（0.18分）之间。

1750~1800年，中国和日本的战争能力都急剧下降。1800年时，清朝拥兵85万，其中25万据说是精锐的满洲旗兵。不过，相对于其庞大数量，我们必须指出其素质、组织和后勤自康熙时代起就在全面衰退。尽管乾隆皇帝因十大军事胜利而在1792年自封了“十全老人”的光荣称号，但实际上他的军队在缅甸、越南和尼泊尔都蒙受了惨重损失。

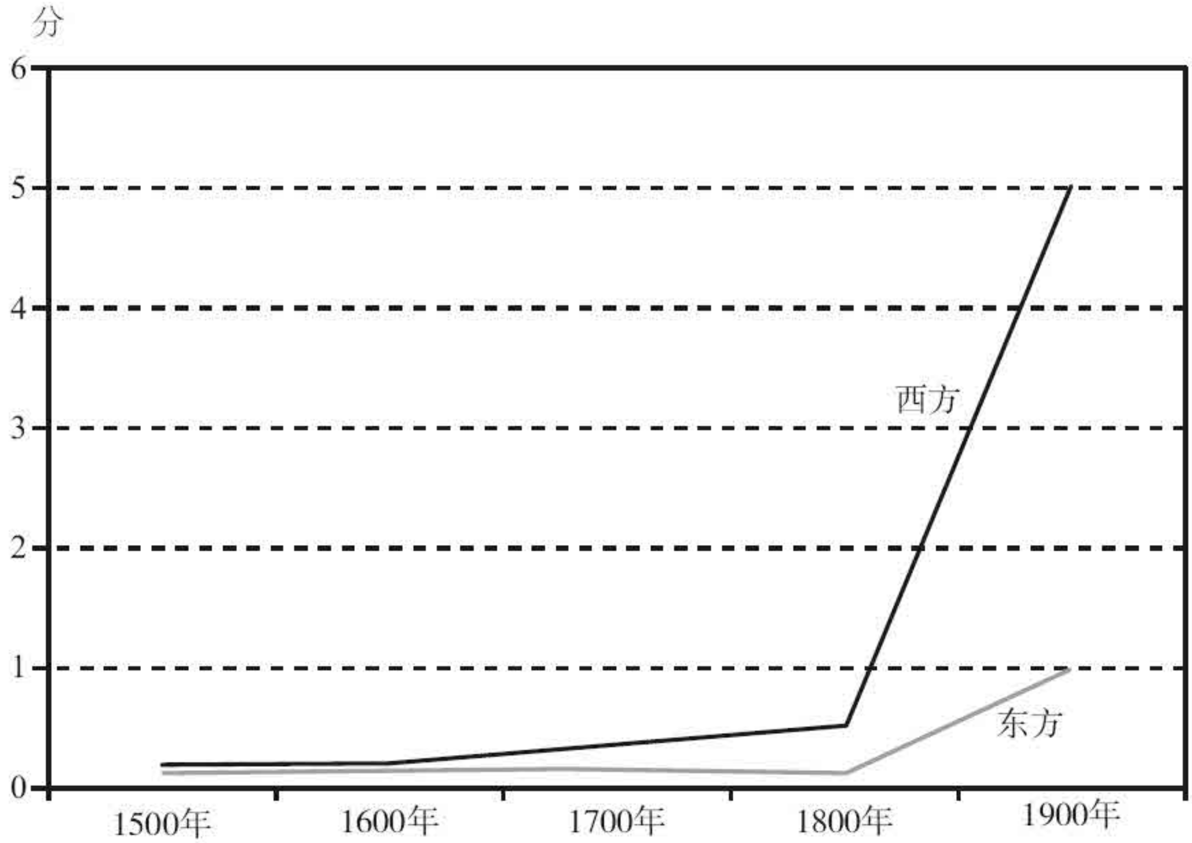
当罗伯特·乔斯林勋爵（我们在本章开头曾引用过他的话）于1840年观看中国陆军和水师的行动时，东方和西方在武器和组织方面的差距已经非常巨大了。英国军官阿尔明·芒廷（Armine Mountain）在一个

著名的比喻中，称中国军队就像是14世纪傅华萨（Froissart）所著《英法百年战争编年史》中的插图，“恰好是他的老版画的主题所呈现出的生活、物质和色彩，不知不觉地跨越了好几个世纪，和所有现代的习俗、发明和改进一起在我眼前浮动。”

我估计西方的战争能力从1300年的0.1分增至1400年的0.11分。如果这个估计是可靠的，而且阿尔明的判断也是可靠的，那就意味着在1700~1840年，东方的战争能力从0.15分降到了大约0.11分。我猜测阿尔明夸大事实的成分极小，于是在1700~1800年，在欧洲人的战争能力增长了将近50%，从0.35分增至0.5分的同时，中国的战争能力却下降了25%，从0.15分降至只有0.12分（日本的军事实力则降得更低）。

这意味着1800年清朝军队的战斗力还不及1600年之前不久与丰臣秀吉作战的明朝军队，不过多少比在克雷西战役、普瓦捷战役和阿金库尔战役中厮杀的骑士和弓箭手们强一些。这也意味着东方的战争能力在19世纪增长了10倍，才使得日本在1900年得到了1分（见图5.15）。

图5.15 1500~1900年，军事革命时代的东西方战争能力



中华帝国和游牧时代（公元前200~公元1500年）

中国在2 000年帝制历史上的大部分时间内，战争能力都强于东方（甚至世界上）的任何对手，但也有例外的时候。其中最引人注目的例外，与我在《西方将主宰多久》一书中称为“游牧异族”的民族有关。

总体而言，我用于衡量社会发展水平的4个特点的分数，都显示出存在相当大的冗余，然而有非同寻常的社会形态来抵消这一趋势。草原游牧民族就是其中最为重要的：这些群体通常在社会组织和信息技术方面都得分极低，在能量获取方面也乏善可陈，但在火药时代来临之前，只有最为高效的农业帝国才能在战场上胜过他们。

在东方，农业帝国显然在大约公元前100~公元100年达到了这一水平。当时汉朝军队经常能击败匈奴军队，然后在7世纪，唐朝军队在与

突厥对峙时甚至取得了更大的优势。然而，直到1700年后，随着火器得到了巨大改善，清朝军队才真正控制了草原。在中原王朝占优势的这些时期之前和之间——即大约公元前200年、200~500年和800~1500年——草原游牧民族的军事实力强于任何农业国家。

在这段漫长的历史时期中，最强悍的草原社会在社会发展指数的战争能力方面，也许平均能得到大约0.1分（浮动为±25%~50%）。最低分（也许是大约0.06分或0.07分）出现在公元头两个世纪，当时罗马、波斯和汉王朝都成功地压制了新的游牧大帝国在草原上的崛起，最高分（可能为大约1.3分，大致是匈奴的两倍）则出现在成吉思汗时代。

这意味着成吉思汗的蒙古游牧部落的武力可能已经超过了罗马帝国，甚至可能与1500年左右的奥斯曼帝国不相上下了。当然，已经无法证明这样的看法是否正确了，但帖木儿的确曾在1402年轻取了奥斯曼军队，并认为他的游牧部落的实力足以在1405年推翻明王朝。另一支蒙古军队的确于1450年俘虏了明朝皇帝，假如他们愿意的话，也很可能攻陷北京城。不论有多少参考价值，“历史上的大战役”游戏也假设成吉思汗的军队在战术上优于恺撒的军队。

然而，尽人皆知的是，游牧民族的统治者在试图将其战争能力转化为政治权力时，都极其困难。只有拥有“半游牧”背景的民族才有可能成功建立统治王朝，如6世纪时鲜卑人建立了隋唐，12世纪时女真人建立了金朝，17世纪时满族人建立了清朝。完全游牧背景的征服者，如13~14世纪的蒙古人，似乎在进行统治一个农业帝国所必需的文化调整时感到太困难。因此，对于整个公元前200~公元1800年，我只给中国的战争能力打分，而不计游牧民族。

东方军事史上这段时期的大致轮廓还是相当清楚的，尽管精确打分又是一个主观的问题。我将从15世纪说起，回溯至公元前200年。

公元1400年，就在郑和远航和永乐大帝征伐草原前夕，明朝的军事力量是非常强大的。皇帝号称统率着一支有3 500条船的近海船队（其中有1 750条战船、1 350条巡逻船和400条武装运输船）和一支120万人的陆军。但实际上，军队的规模要小得多，不过在1414年对草原最大的一次征伐的确投入了大约50万兵力。明朝的战争能力无疑比当时的奥斯曼帝国要强（我给奥斯曼帝国打的分数为0.11分），但恐怕不及13世纪中期蒙古人鼎盛时期，因此，我为1400年的东方估计的分数是0.12分。

到1300年时，蒙古人在夺取了全中国后，军事实力比13世纪中期巅峰时可能有所下降，但依前现代的标准依然令人生畏。蒙古人建立的元朝甚至重建了中国水师，修复了自13世纪中期就处于失修状态的战船，据说还在1274年派遣了4 500艘战船，搭载15万士兵远征日本。我为之打0.11分，稍稍低于明朝早期全盛时期，但这只是猜测；估计为0.1分或0.12分也是合理的。

尽管宋朝有不尚武的名声，但在10世纪晚期，其军队仍在迅速发展。据说997年太宗驾崩时，宋军有65万之众；1022年仁宗时，兵力接近100万。11世纪的王安石变法引发了财政危机，使得领军饷的职业士兵逐渐减少而民兵开始增多，削弱了整体力量，但军队依然强大。1081年时，正规军有32万人，还有同样数量的勤杂人员，得到庞大的中央军械库的支持。甚至在12世纪二三十年代，朝廷仍然能向战场上投入10万~20万兵力。

12世纪的南宋统治者为了与女真族作战，大大加强了水师，制造了更大的战船，包括能抵御风浪的明轮船，还研制了许多新武器，如弩炮、火箭和喷火器等。12世纪的明轮船有可能长达60~90米，配有8个轮子，可载700~800名水手。到12世纪30年代时，最大的船的长度超过了100米。到1200年时，一些船甚至装上了铁板，做装甲用。

1000~1200年，宋朝的军事力量显然比当时四分五裂的西方要强大得多。在西方，1000年时军事力量最强的可能是拜占庭帝国，1100~1200年是塞尔柱突厥。我给1000年的拜占庭打0.06分，给1100~1200年的塞尔柱人打0.07分和0.08分。我姑且建议给1000年东方的战争能力打0.08分，给1100~1200年的东方打0.09分。这意味着宋朝的战争能力即使在其巅峰时期，也不及罗马帝国。

然而，唐朝的战争能力更为接近罗马帝国。8世纪早期的文献显示，唐朝有大约50万军队，军队高度集权，有严格的纪律，有长期服役的专业化部队。

唐朝的军事力量依赖于将草原的重型骑兵与6世纪北魏、北周及其继承者隋朝发展起来的大规模步兵结合了起来。这一过程始于5世纪早期鲜卑征服中国北方大部，5世纪晚期，北魏孝文帝改革加速了其进程，但即使到了6世纪30年代，10万人仍被认为是一支庞大的军队。

直到6世纪晚期，随着中国北方的巩固，才产生了更加强大的军事力量。589年，隋文帝征服南方时，曾在长江流域屯兵518 000之众，并得到有5层甲板的战船的支援。这种战船可载800人，装备有尖刺状的伸臂，可固定住敌方的战船，并供本方士兵登上敌船（令人惊讶的是，听上去很像罗马人在公元前3世纪60年代晚期制造的带有乌鸦吊桥的五桨木船）。

在隋朝建立了庞大的水军，于589年统一中国，于612~614年与高句丽进行了灾难性的战争后，唐朝水军逐步缩小。然而，保留的规模依然算是很大，因为没有国家能从海上对该帝国形成实质性的威胁，（至少到755年前）中国的国内和平也不要求江河上有庞大的武装存在。只有极少数的情况下需要船只，当7世纪60年代与朝鲜的战争重启后，强大的唐帝国也能够在短时间内建造或征用数以百计的船只，发动大规模战役。

755~763年的安史之乱极大地削弱了唐帝国。763年11月，一支吐蕃军队竟然攻破了长安城，而且在接下去的两个世纪里，中国的军事能量都消耗在连绵不断的内战和算是部分成功地抵御吐蕃侵袭的战争中。在中央权威崩溃后，各藩镇还保持着自己的军队，但即使最大的藩镇（例如平卢军）的军队人数也不到10万。而且藩镇的军队普遍装备、补给和统领都很差。

在5世纪早期鲜卑统一中国北方之前，军队规模都相当大，但都远不及唐朝时期战斗力强。例如，公元279年，晋朝召集20万大军和支援战船征讨长江以南地区。这场战争与隋文帝于589年在同一地区为同样目标而发动的战争非常相似，但晋军的人数只是隋军的40%，且组织水平更接近于汉朝而不及隋朝。

在接下去的200年里，骑兵主宰了中国战场。出土文物、陶俑和墓中浮雕提供了大量关于武器的信息，显示4世纪时骑兵普遍使用了马镫。很多坟墓中出土的物证都表明冲击武器和马的护甲越来越普遍地得到应用，说明战术发生了重大变化。

虽然200~400年的战争能力显然有重大进步，给这段分裂时期打分却仍然不容易。东方的军事力量从来没有衰落到西方7~9世纪的羸弱水平，国家的基础设施始终得以保存。即使在4世纪，5万~10万人的军队也比比皆是。尽管攻城器械基本上从北方军队中消失了，但南方的堡垒依然坚固。400年后，军事实力增长得比以前更快。从我给600年的隋朝军队打0.09分回溯，我给500年打0.08分，给200~400年这段时期打0.07分。

变化极小的分数遮掩了重大的变革，但除非我们将公元600年的分数推得更高，我们就不得不或者（像我那样）设想公元200~400年的变化没有大到能在指数上显示出来，或者认为分数在公元200年后的某个时候降到了0.07分以下，然后又迅速攀升了。0.06分将相当于西方在6

世纪早期或大约公元1000年时的水平，但我从文学作品中得到的印象是，在整个分裂时期，东方的战争能力仍然高于那些水平。

西汉王朝（公元前206~公元9年）的战争能力还要更高。公元前3世纪的大战，催生了交战双方都动辄上万人的大规模步兵部队。他们娴熟地运用围城技巧和后勤保障手段，还开创了大量的军事理论。

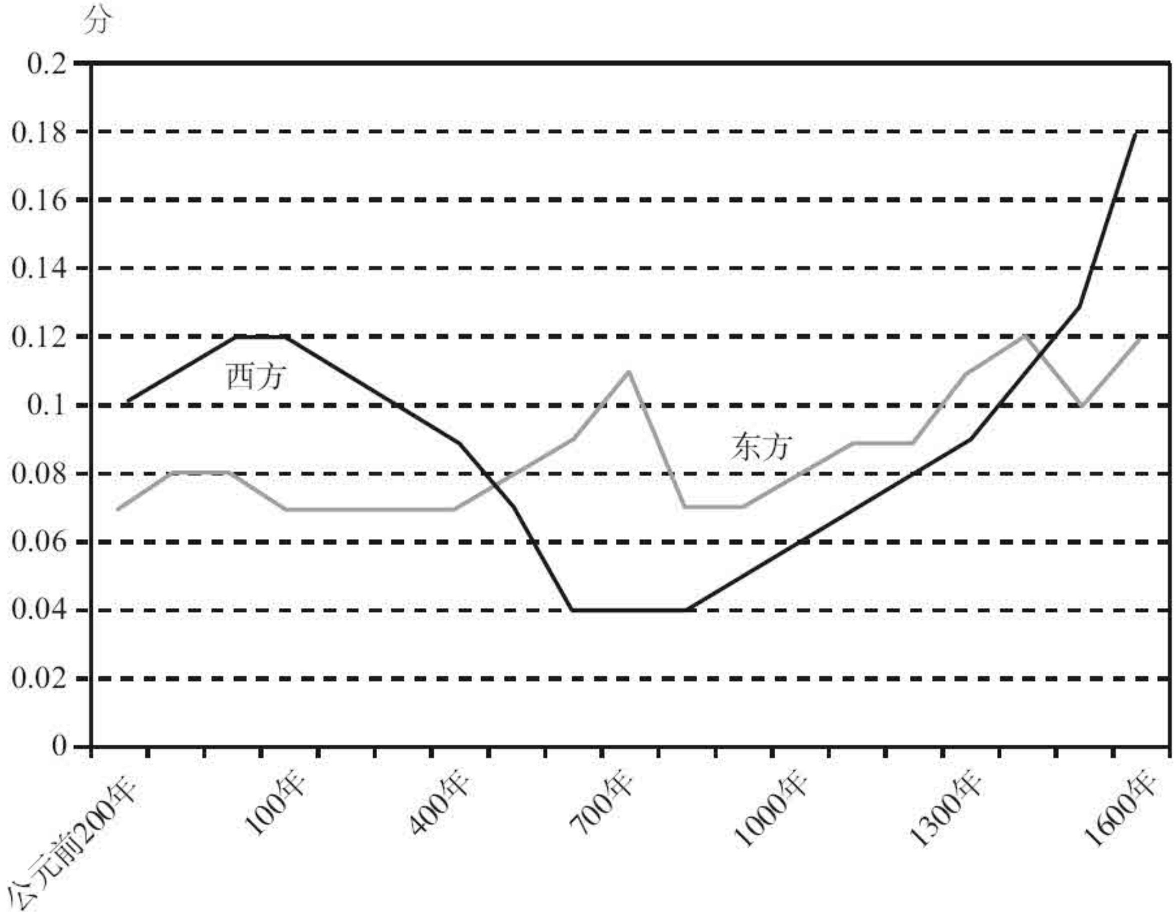
公元前200年时水军还很弱，因为控制江河湖海没有什么重要意义；骑兵也是如此，而且许多军队都还在使用青铜武器而不是铁制武器。然而，在接下去的两个世纪里，铁制武器逐步取代了青铜武器，随着主要战争从中国军队之间的内战转变为抵御匈奴游牧部落的战斗，骑兵的重要性也在不断增长。

西汉军队的规模是起伏不定的，公元前200年后随着皇帝削藩，军队人数有所下降，但为了打大仗，此后又有所攀升。如公元前97年武帝出击匈奴时，就派出了14万步兵和7万骑兵。不过总体趋势是下降的，东汉王朝（公元25~220年）于公元31年废除了普遍征兵制，并认真开始在帝国核心地带实行非军事化。到公元1世纪50年代时，汉帝国的军事制度改为将大量军队（通常都是处于帝国极松弛的控制下的盟国骑兵）驻扎在边境地区，帝国核心地带只留下一支规模极小的常备军，大约只有4万人。

汉朝军队的战斗力似乎始终没有达到罗马帝国军队的水平。我为公元前100~前/公元1年的汉军打0.08分（罗马巅峰时期为0.12分），然后公元100~200年稍降一点，为0.07分。对这段时期的大部分估计都带有强烈的主观色彩。在当前的证据条件下，即使将西汉的得分提高到0.1分，也不为过。不过，除非我们认为汉朝的军事实力实际上与罗马势均力敌，否则不可能将东方的战争能力分数改变到足以严重影响社会发展指数的程度。

公元前200~公元1500年西方和东方战争能力曲线（见图5.16）的形状显示，在火药革命爆发之前，恐怕任何国家（包括最强大的草原联盟）都不可能使其军事力量超越0.1~0.12分。尽管进行跨越漫长时间和广袤空间的广泛对比极其困难，但铁制兵器武装下的罗马帝国军队、唐朝初期军队和明朝初期军队的战斗技能，似乎的确是达到了大致相同的水平。较而言之改进指挥与控制、加强后勤的空间都不大了，因此也不可能比这些时段的水平再高很多了。

图5.16 公元前200~公元1600年东方和西方的战争能力



中国早期（公元前1600~前200年）

正如对西方早期战争能力的分析一样，在指数中可能出现的极小分数只能对图解产生影响（见图5.13）。

考古学家发现了大量中国史前时代使用武力的证据，但直到公元前第二个千年早期，我们才看到金属武器的经常使用，以及我们可以合理地视为国家级战争的军事组织的迹象。依照我对西方早期战争能力采用的原则，我将最低的分数0.01分给公元前1600年，这与通常被认为与商朝的到来相关的一些事件是大致相符的。公元前第二个千年中期东方的战争能力，与公元前第三个千年的西方核心地带的大致处于同一水平。战争中通常使用青铜武器，参战人数只有几千名，都是没有盔甲的民兵，没有骑兵或战车，也没有专门为打仗建造的船，只有非常简单的防御工事。

东方和西方战争能力的时间滞差在公元前第二个千年的晚期急速缩小了。这大概与战车从中亚向这两个地区的推广（大约公元前1800年传至西方，公元前1200年传至东方）大有关系。商朝晚期的战争似乎比商朝早期规模要大得多。按照甲骨文上的标示，商朝的远征军通常大约有3 000多人，但至少有一次，商王武丁和夫人妇好召集了10 000人。到公元前1200年时，这些军队已经用上了战车，但似乎局限于主要将战车用于军官的车乘。公元前1200年时战争能力的增强足以将分数提至0.02分。我估计西方的分数是在公元前1800年时升到了0.02分，说明东方和西方军事力量的差距缩小到了大约600年。

在西方，军事变革的步伐在公元前1500年后加快了。在地中海东部地区的王国中，战车部队成为主力，在高度集中的统领下，越来越多的专业化部队建立了起来。西方战争能力的得分在公元前1400年增至0.03分，公元前1300年增至0.04分。在公元前第二个千年的晚期，东方似乎也经历了一个类似的军事能力加速增强的时期。到公元前1000年时，周朝也大规模地使用起战车来，就像自公元前1500年起的西方军队。据（公认是将近1 000年后编撰的）《史记》记载，公元前1045年，周武王在推翻商朝的战争中，统率了45 000名步兵、6 000名盟军和300辆战车。即使实际数字只有一半，这支军队按照西方公元前1500年的标准，也相当可观了（尽管不会给公元前13世纪的西方国王们留

下什么印象）。因此我提议给公元前1000年的东方战争能力打0.03分。

周朝军队在公元前10世纪似乎更加壮大了，并使王权跨过了渭河，扩展至黄河流域。然而，在公元前957年周昭王在汉水流域遭遇惨败后，国家开始分裂。给这样一个鲜为人知的体系打分无疑是武断的，但公元前1000~前900年周朝军力的增长不足以使得分升至0.03分以上，但公元前850年后国家陷入的混乱却足以使得分降至0.02分。

军事力量在公元前第一个千年的中期和晚期迅速而稳步增长。最重要的变化在组织方面，从贵族征税征兵，并坐在战车里指挥自己的部队，转为统治者征税，征召自由的农民组成大规模的步兵部队。公元前7世纪时，10 000人的军队仍被视为庞大。然而，到了公元前6世纪晚期，主要的军队能召集到50 000人，而一个世纪后，最大的军队的人数还会再翻上一番。

公元前5世纪时，中国学者以能投入战斗的战车数量来衡量国家的大小：千乘之国（1 000辆战车可能只是公元前632年晋文公投入城濮之战的战车数量的一半）为小国，万乘才是大国。然而，公元前4世纪和3世纪，军队的规模呈爆炸式增长。我们的文献资料提供的数据（达到了60万）经常受到怀疑，但到公元前250年时，秦国无疑能一次性投入数十万兵力。铁制兵器直到公元前200年后才普遍使用，但公元前4世纪和公元前3世纪，我们也看到了战车被骑兵取代，围攻战也有了很大的进步，中国北方许多地区都建起了泥砖长城，以抵御草原袭击者。

在图5.13中，我标绘的东方战争能力的分数在公元前700~前100年是稳步增长的，从0.02分增至0.08分。这无疑是过分简单化了，增长率在公元前400年后很可能加速了，但鉴于涉及的分数实在太小，目前这样标绘，比插入停滞时期和骤变时期，随意性似乎要小一些。

工业革命改造枪炮

跟社会组织一样，战争能力也是能量获取的功能之一。能量获取水平稍微提高一点点，通常就能引发战争能力的巨大增长（见图5.17和图5.18）。

图5.17 公元前14000~公元2000年西方能量获取水平与战争能力的对比（以对数－线性标尺标绘，以社会发展分数衡量）

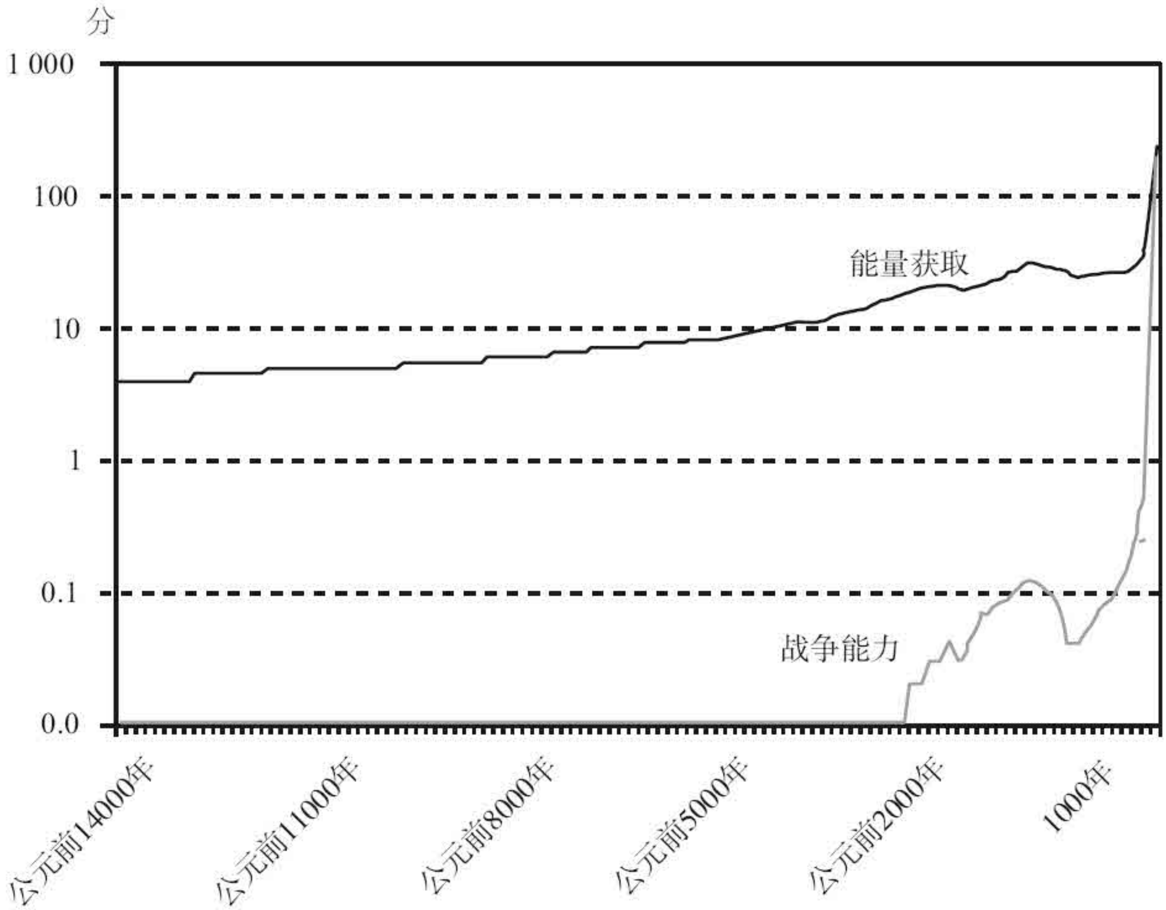
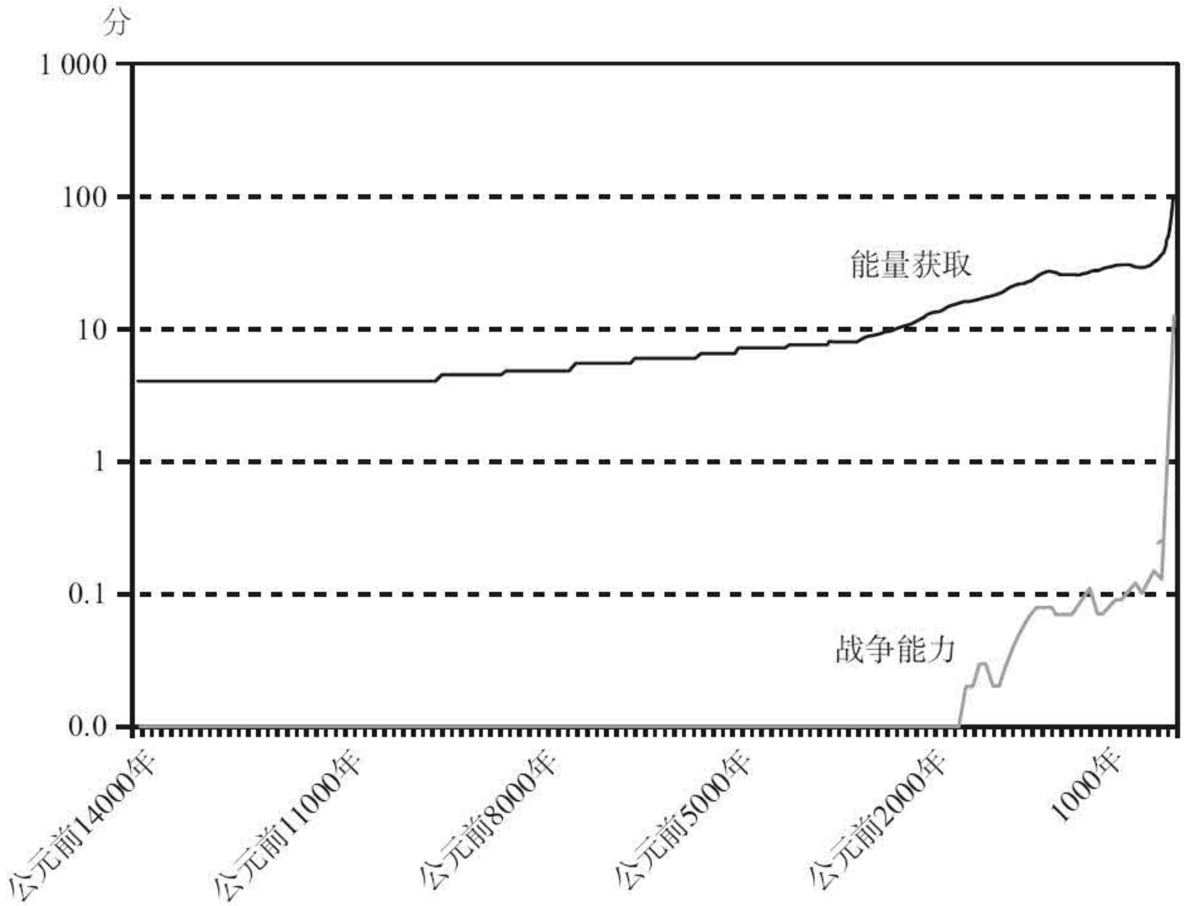


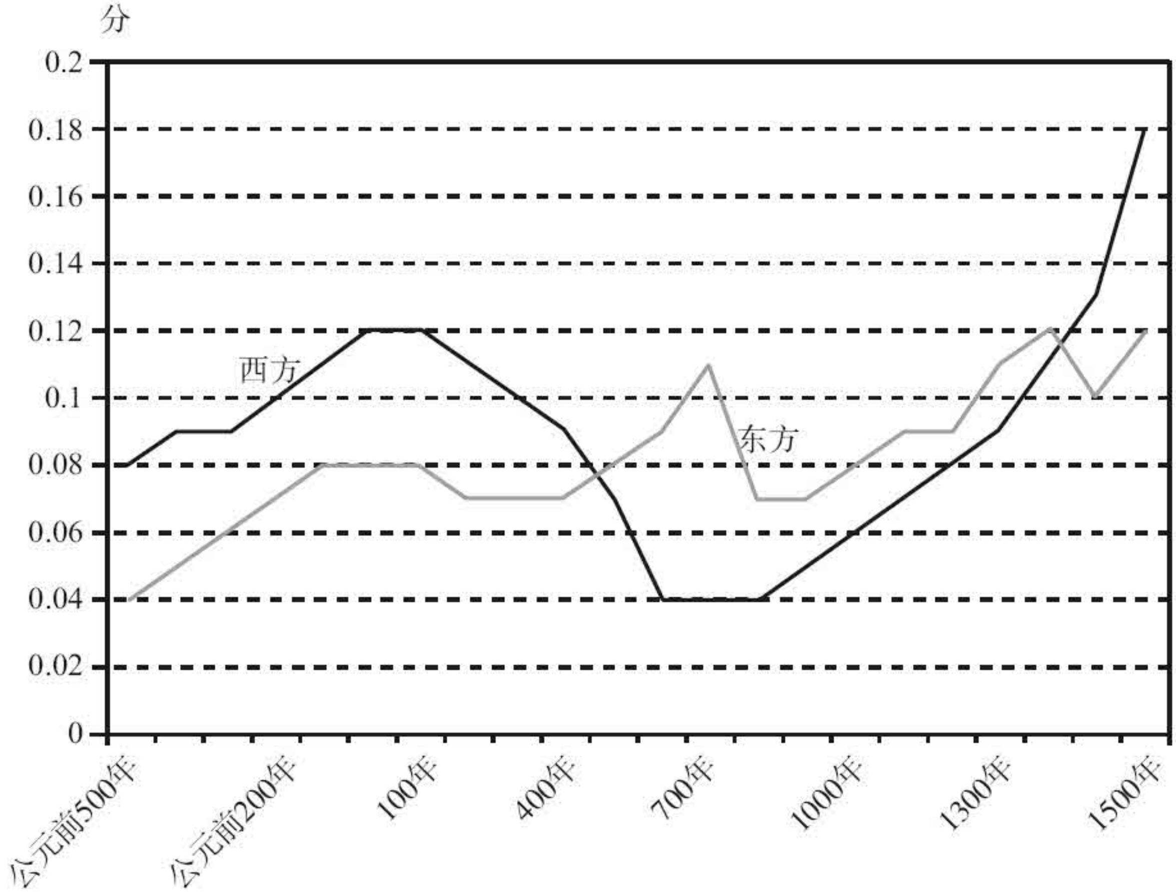
图5.18 公元前14000~公元2000年东方能量获取水平与战争能力的对比（以对数－线性标尺标绘，以社会发展分数衡量）



东方和西方一样，战争能力在经过了一段增长缓慢得在发展指数上难以衡量的漫长时期后，又突飞猛进了一段，在大约1 000年间（西方为公元前1800~前500年，东方为公元前1200~前100年）从0.01分跃升至0.08分。继而在接下来的将近2 000年间，东西方的战争能力又都似乎在不断地冲击着我们也许可称之为“硬天花板”的上限，即0.08~0.12分（见图5.19）。

在0.08~0.12分这个范围内，东方和西方的战争能力经历了多少不同的道路。在东方，我们看到了一系列的波峰（公元前100~公元100年、公元700年、公元1400年）被崩溃（公元200~400年、800~900年）分开。每个波峰都比前一个要高（公元前100~公元100年为0.08分，公元700年为0.11分，公元1400年为0.12分，在大约1500年经历了一次轻微的下跌后，于1600年又恢复到这个水平）。

图5.19 公元前500~公元1600年，军事上的“硬天花板”



前现代时期的西方战争能力经历了一段更富戏剧性的历史，从公元前500~前100年（如前所述，图中公元前4世纪那段时间的线是平直的，只是为了将极小的分数凑整）相当平稳地增长到一个非常高的水平——0.12分，然后又一路下跌，到公元600年时降至0.04分，公元800年又开始平稳上升，约于15世纪40年代突破了0.12分的上限——此时正是从奥斯曼土耳其前往勃艮第、法兰西和英格兰的枪手们开始实验新武器和新战术的时候。

图5.19恰好说明了早期现代欧洲军事革命理论家们的观点：枪是欧洲军事腾飞的必要条件，但不是充分条件。是枪使得15世纪时的欧洲军队在很多重要方面超越罗马帝国成为可能，然而是战术、后勤、

指挥和控制等方面都发生了变化，才使得新武器的潜力得到了充分发挥。

正如肯尼思·蔡斯（Kenneth Chase）所指出的，这些变革之所以始于欧洲，尤其是西欧，而不是中国、印度、伊朗或土耳其，是因为：（1）欧洲距草原较远，这使其难于维持较大规模的骑兵部队，意味着其战场上保留着大量运动缓慢的步兵，因此发射较慢的枪仍是有效的；（2）欧洲有大量筑有围墙的城市，使得大炮有较大的用途；（3）欧洲在政治上四分五裂，意味着战事不断，因而革新是有益的。

表5.2 约1400~1700年促成军事革命的因素

	步兵	城市	分裂程度
基督教世界	非常重要	非常重要	非常重要
伊斯兰世界	比较重要	不重要	比较重要
东方	不重要	比较重要	不重要

表5.2粗略总结了地区差异，将欧亚大陆幸运纬度带上的地区分为3部分，分别是基督教世界、伊斯兰世界和东方。在伊斯兰世界（大致从巴尔干半岛到印度），步兵经常在战场上发挥重要作用，但草原骑兵通常起主宰作用，很少采用战术，因为并不能决定战争的胜负。虽然印度在政治上有些四分五裂，奥斯曼帝国与波斯、埃及及其他邻国都进行过长期的战争，但政治争端并不像欧洲那样剧烈。东方则相反，壁垒森严的城市是首要的军事目标，但除了1592~1598年在朝鲜发生过大战外，鲜少发生国家间的战争，主要的战争是在中国的边防军与草原骑兵之间进行的。中世纪晚期和现代早期，欧洲发生军事改革的诱因要比欧亚大陆任何其他地方都要大得多，所以是西欧人最先开发了火枪和火炮的潜能，并为进行更现代化的战争而组织社会。

这就是说，组织效率不可能将现代早期的战争能力在社会发展指数上再提高哪怕1.0分。只有工业革命和将自然科学手段运用于运输、

供给和火力，才能做到这点。将图5.17和图5.18，与图4.2和图4.3对比，就可看出工业革命对战争能力的影响比对城市规模要大得多。

---

1. 1节=1海里/小时≈1.852公里/小时。——编者注
2. 1磅≈0.45千克。——编者注

The Measure of Civilization

How Social Development Decides the Fate of Nations

第六章

信息技术：它影响了社会

读、写和计算

除了极少数例外之外，人类与所有其他动物不同的是，能够随时间推移，通过积累信息、观念和优良传统来实现文化进化。早在180万年前的直立人时期，原始人类也许就有了与现代语言相似的语言，而海德堡人——尼安德特人和现代人类的共同始祖——有可以发出语音的舌骨和也许能分辨出交谈语言的发音的内耳。然而，过去15万年来智人的进化，才代表着这方面的革命。

几十万年来，信息的传输和储存完全依赖于语言和记忆。最早的通过实体信息符号进行交流的确凿无疑的证据，要回溯至将近10万年前，南非克拉西河1号洞刻有符号的赭石碎片。然而，这类符号不仅极少，也非常简单，直到大约5万年前，才在所有发现人类的地方突然普遍起来（至少是依照史前的标准化）。考古学家们通常称之为“人类意识的大爆炸”。

最早的明确表示数字和话语的符号，于5 000年前较常出现于亚洲西南部，自那以后这些技术就传遍了全世界。几乎是理所当然地，我们相对非常清楚自文字和数字产生后的信息技术的历史，因为每一件从古代留存下来的文件，本质上都是这样的技术完善和推广的例证。因此，我们能够比较详细地探查到存储和交流信息的系统是怎样产生的，获取信息是怎样变得越来越容易的，以及各种各样的技术是怎样不断完善的。

存储和传播信息的能力对于掌握智力环境至关重要，并且本身就是社会发展这个概念的一个基本的组成部分。然而，尽管人类写作和计算的自然痕迹有考古证据，但衡量不同技术的应用程度仍是非常困难的。欧洲的一些历史学家曾勇敢地尝试估计过去两三千年有多少人

能读、会写以及语言水平。计算能力尽管明显更重要，但得到的关注度不及读写能力，但还是有一些极富价值的研究成果。

自20世纪80年代起，语言学家中就有反对量化的声音，许多欧洲学者认为，欧洲有那么多种类的语言，计算有多少人能读、会写是没有意义的。这种认识的前半部分（欧洲有许多种类的语言，因此也就有同样多种类的计算方法）无疑是正确的，但后半部分（计算有多少人具有读写能力是无意义的）并非必然。只要我们清楚读写和计算的意义是什么，并且承认其他历史学家在问其他问题时也许会选择不同的定义，那么量化就仍然是必要的手段。至于在人文科学和许多社会科学领域更广泛的反对量化之声，那在很大程度上都是对于要回答什么问题的质疑之声。

将信息技术作为社会发展指数的一个特性，需要分别计算下列分数：（1）在特定时间点上，技术在东方和西方的成熟程度；（2）技术的使用范围。然后将两个数字相乘，以得出东方和西方信息技术在历史上的一系列分数。

跟考据战争能力的情况一样，最大的困难不在于前现代时期的证据稀少，而在于20世纪技术成熟度的巨大飞跃，使得公元2000年的信息技术很难与先前时期对比。我在《西方将主宰多久》一书中曾指出，摩尔定律称信息存储和检索的成本效益自1950年以来每18个月翻一番，也许意味着公元2000年西方信息技术的分数比1950年高出10亿倍还多。公元2000年西方的分数是250分，实际上我们回溯不到1970年时，就已经跌至可衡量的最低分数0.01分了。

我们很多人还记得20世纪70年代的卷盘式磁带机和有大型主机的计算机，与我们这个文明时代的音乐播放器iPod和平板电脑iPad等相比，那些简直像是古代的机器；然而假如你认为在人类首次登上月球的那个时代的信息技术太过原始而无法衡量，那也是很荒谬的。计算信息技术的分数要求对不同种类的系统加权，并确认它们之间的变化

不是线性的或直接的。文字并没有替代语言，电话或推特也没有替代面对面的交谈。新的信息技术形式也许最终能完全替代过去几十万年来发展的形式，但这种情况还没有发生，而且在计算信息技术的历史分数时，我们还必须认清复杂的重叠模式。

有多少人会读、会写、会计算，他们的技能处于什么水平，他们使用什么信息技术，其证据是零散破碎的，也会引发很多不同的解释；而且还要考虑到时代变迁中不同的人对不同的变革是各有偏好的，这又增加了计算中的主观性。因此信息技术的分数比另外3项特性争议更多。

信息技术记分牌

给信息技术分类的困难，要求在计分时采用两段法。

1.技能。遵照历史学家们的普遍范例，我将作为研究对象的人群，按照使用他们那个时代的信息技术时的能力，划分为3个技能水平：初步、中等和充分。同样遵照标准范例，我在划分每个类别时都将门槛设得较低。“初步”包括能够读和写名字，或者记录简单的数字；“中等”意味着能够读或写简单的句子，或者用加、减、乘、除来解决基本的问题；“充分”意味着能够读或写更连贯的文章，或使用更高级的数学技巧。

一些人类学家和历史学家曾提出，这样的定义是以欧洲为中心的，因为在有些文化传统中，语言和数学是以完全不同的方式运作的。然而，尽管这个问题值得进行更多的研究，但目前似乎还没有什么实证研究支持这样的主张。例如，将读写能力划分为初步、中等和充分3级水平，是中国在1950年开展扫盲运动时独立创制的，规定充分

的读写能力为认识1 000个以上的字，中等文化水平为认识500~1 000个字，而基本的读写能力为认识300~500个字。

利用一切可利用的研究成果（如有可利用的专家的定量估计则直接利用，如无则通过专家的定性讨论进行推算），我将不同时期人口中的成年男子按照充分、中等和初步3个类别进行了划分。如果人口中每1%的成年男子能够进入充分技能类，则可得0.5 ITP（信息技术分）；每1%的成年男子能进中等技能类，可得0.25 ITP；每1%的成年男子能进初步技能类，可得0.15 ITP。

这些数字是且只能是掌握信息技术的不同水平之间的差异的武断估计。这也许对于某些时间和地点是非常合理的，但对于另一些则大谬不然。然而，打分时坚持一贯性似乎比虚伪而高度主观地追求更加精确更为重要。将分数相加起来，会得出每个时期唯一的男性ITP。如果认为我为初步、中等和充分技能类别提出的数字不合理，批评家们当然可以用其他数字做实验，然后看看需要多少数字才能改变结果，并且对社会发展指数产生重大影响。

关于女性读写和计算能力的证据普遍比男性要少，不过我们可以肯定在20世纪前的大部分时间和地点，会读、写和进行数学计算的女性要少于男性（通常是要少得多），并且通常都处于较低水平。

史前时代的男女差异无疑没有可靠的统计数字，这意味着我又一次只能瞎猜，唯一所受的约束是从历史资料中得来的总体印象。然而，做出清楚的猜测比留下含糊的假设更具建设性，所以我冒昧地提出一系列估计数字，抛砖引玉。然后我将估计的每个时期的性别乘数运用于男性ITP分数，得出女性的ITP分数；将这两个分数相加，就能产生一个东方和西方在特定时间点上的唯一的ITP分数。

在公元2000年的西方核心地带，我将100%的成年男子划入充分技能类，得到的男性ITP分数为50（即100%×0.5）；而女性的技能分数

也是男性的100%，于是女性的ITP分数也是50分（即男性的50分×100%）。因此，公元2000年西方ITP得分为100。

西方核心地带的职业文字和计算从业者通常设定的初步、中等和充分技能的标准比历史学家要高得多，因此他们不仅不会同意我所主张的100%男性都有充分的技能，也不会同意女性的计算技能与男性相当。然而，尽管为初步、中等和充分的读写技能和计算技能设定极高的标准对那些寻求在复杂的21世纪社会提高标准的人非常适宜，对于长期的跨文化比较却无所助益，因为那将使1900年之前的所有分数都降为零。

2. 技术的变化速度和所达范围。计算分数的第二个阶段，是再设立一个能反映存储和交流信息的技术的变化速度和所达范围的乘数。我将处理信息的工具分为3个更广泛的类别：电子媒介（2000年时在东方和西方广泛应用），电媒介（1900年时在西方广泛应用，但在东方没有应用），还有电之前的媒介（在西方可能使用了11 000年，在东方可能使用了9 000年）。

我为公元2000年西方使用的最先进形式的电子媒介赋予的乘数值为2.5。2000年时东方也使用类似的媒介，但不够普及。电话（既包括固定电话，也包括移动电话）和电视在东西方大致同样普及，但电脑和互联网主机在西方更为普及（美国每100人中有62.3台计算机，中国香港为38.5台计算机，日本为34.9 台计算机；美国每100人中有375.1台互联网主机，中国台湾为97.3台互联网主机，日本为72.7台互联网主机）。由于2000年西方的乘数为2.5，我为东方核心地带使用的乘数为1.89。2000年西方在社会发展指数上的信息技术得分为250分（即100 ITP×2.5）；东方的得分为189（即100 ITP×1.89）。

公元2000年西方核心地带电子乘数定为2.5，是因为为这一特性所设的最大分数是250分，但为电媒介和电之前的媒介设乘数值就要难多了。我不知道以前是否有人尝试计算整个20世纪信息技术的功能的总

体增长量，但根据专家文献，我的猜测是2000年可用的电子媒介比1900年西方可用的电媒介，功能增长了大约50倍。这意味着1900年西方核心地带的乘数为0.05。

19世纪的信息技术也有非凡的进步，不过显然与20世纪无法相比。1900年西方可用的电子媒介功能比1800年时可用的电子之前的媒介增长了5倍，这样便得出1800年的乘数为0.01。我将此作为所有电之前的信息技术系统的基准面，包括大约公元前9000年时出现在西方，公元前6250年时出现在东方的，有记录的最早的用视觉符号所做的实验。

也许有人不同意我提出的数字，我对电之前的信息技术的粗泛分类无疑也存在许多变量。历史学家尤其会注意到我对印刷媒介和印刷产生之前的媒介没有做类别上的区分，尽管印刷机的发明在15世纪对欧洲精英文化、17世纪后对东方精英文化所产生的影响已经众所周知。

我做出这样的决定，是因为印刷的主要贡献是产生更多、更便宜的物质，而不是像电报在19世纪、互联网在20世纪那样改变了信息存储和检索的方式，这些纯粹数量上的变化已经作为因素计入了指数。然而，即使其他学者不同意这种看法，涉及1900年前的信息技术分数的数字也实在太微小——甚至比战争能力的数字还要小——将不会使乘数有太大改变，也不会对最终的社会发展分数造成太大的影响。

出于同样的原因，我也没有区分符号形式，将字母的、章节的、表意的以及其他的文字形式均视为电产生之前的系统的变量。这样过分简化了事实，但是因为：（1）对文字系统相对效率的判断太容易陷入局限于文化的价值判断中；（2）公元1700年前所有时间点上的微小分数意味着任何看似合理的调整都不会产生重大影响，所以我决定索性对于电之前的各种信息技术系统一视同仁，只关注它们的应用范围。

最后，我没有将电之前的计算工具单分成一类，比如大约公元前2500年最早发现于美索不达米亚的算盘，或者其简单形式大致同样古老的印加结绳话。这与我没有区别对待印刷机是同样的原因：电之前的计算工具加快了计算速度，改善了精确性，但并没有像计算机那样改变计算过程。

图6.1以线性–线性标尺展示了我计算的分数：1900年西方的分数不过是勉强看得到，但更早时期的分数就都看不到了。图6.2在对数–线性标尺上展现了同样的数据。不同在于，将西方1500~1800年的乘数改为0.02以反映印刷机带来的重大影响，将东方1400~1900年的乘数改为0.02以反映这段时期印刷术的巨大扩张，在线性–线性图（见图6.3）上造成不了可见的变化，在对数–线性图（见图6.4）上只能造成极小的变化。

图6.1 公元前4000~公元2000年东方和西方的信息技术（以线性－线性标尺标绘）

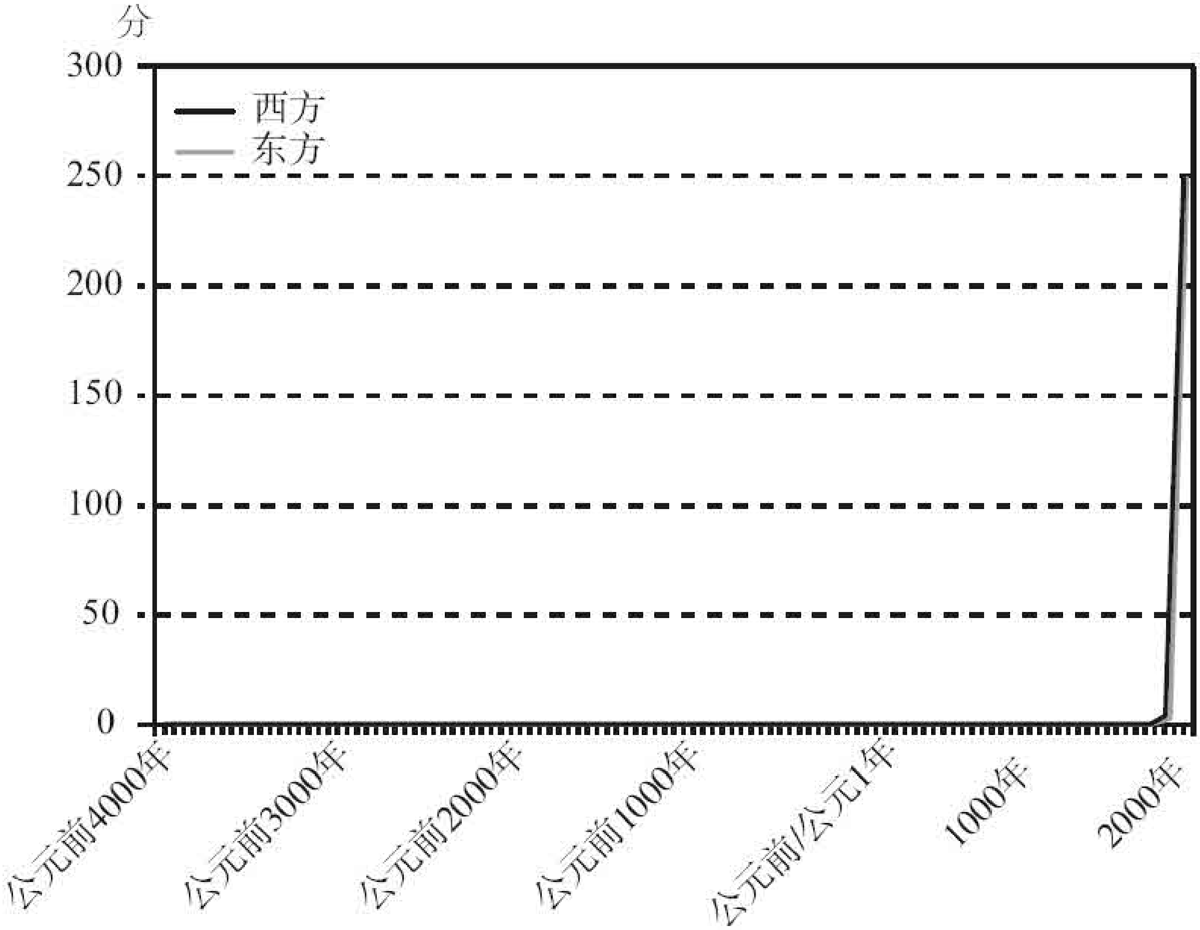


图6.2 公元前4000~公元2000年东方和西方的信息技术（以对数－线性标尺标绘）

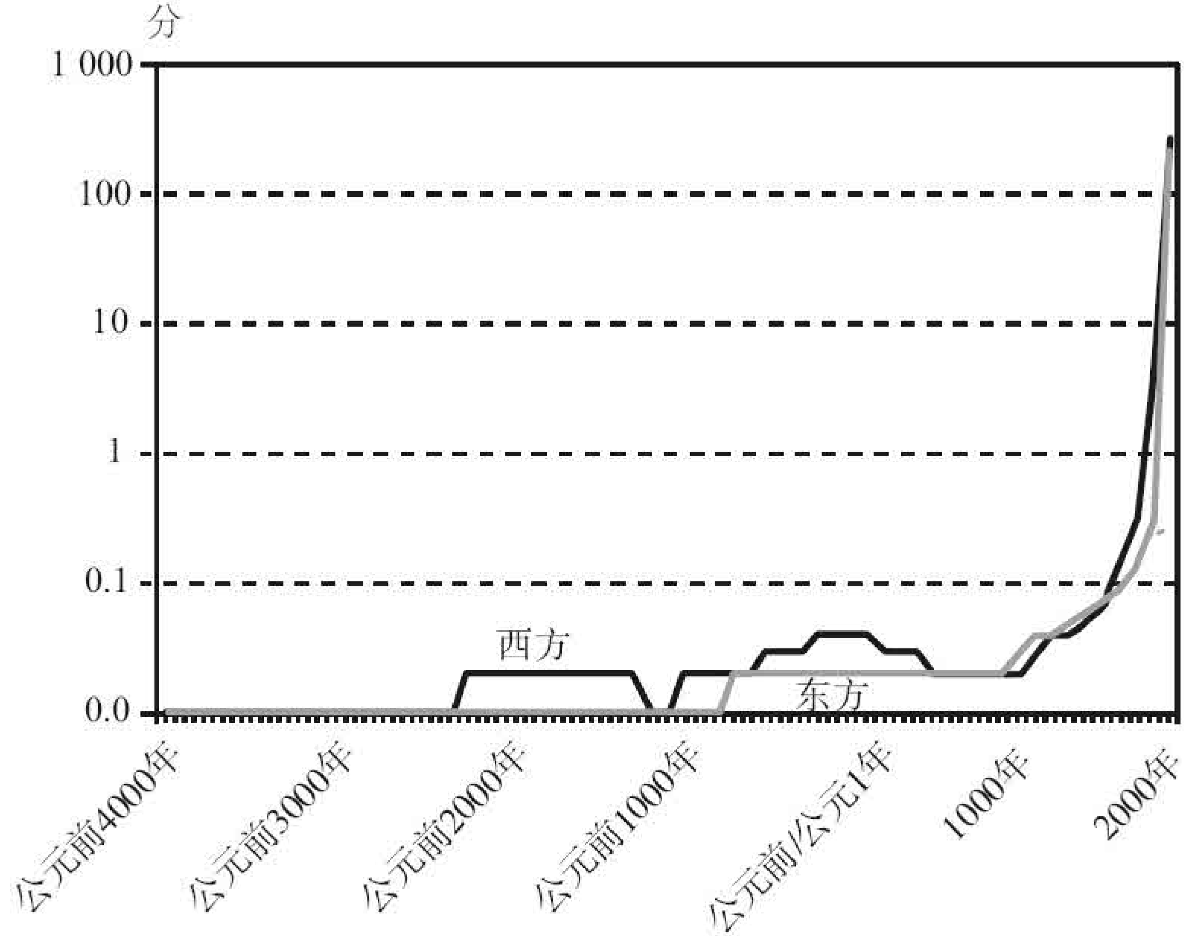


图6.3 公元前4000~公元2000年东方和西方的信息技术（东方1400~1900年，西方1500~1800年因印刷术而分数翻番）

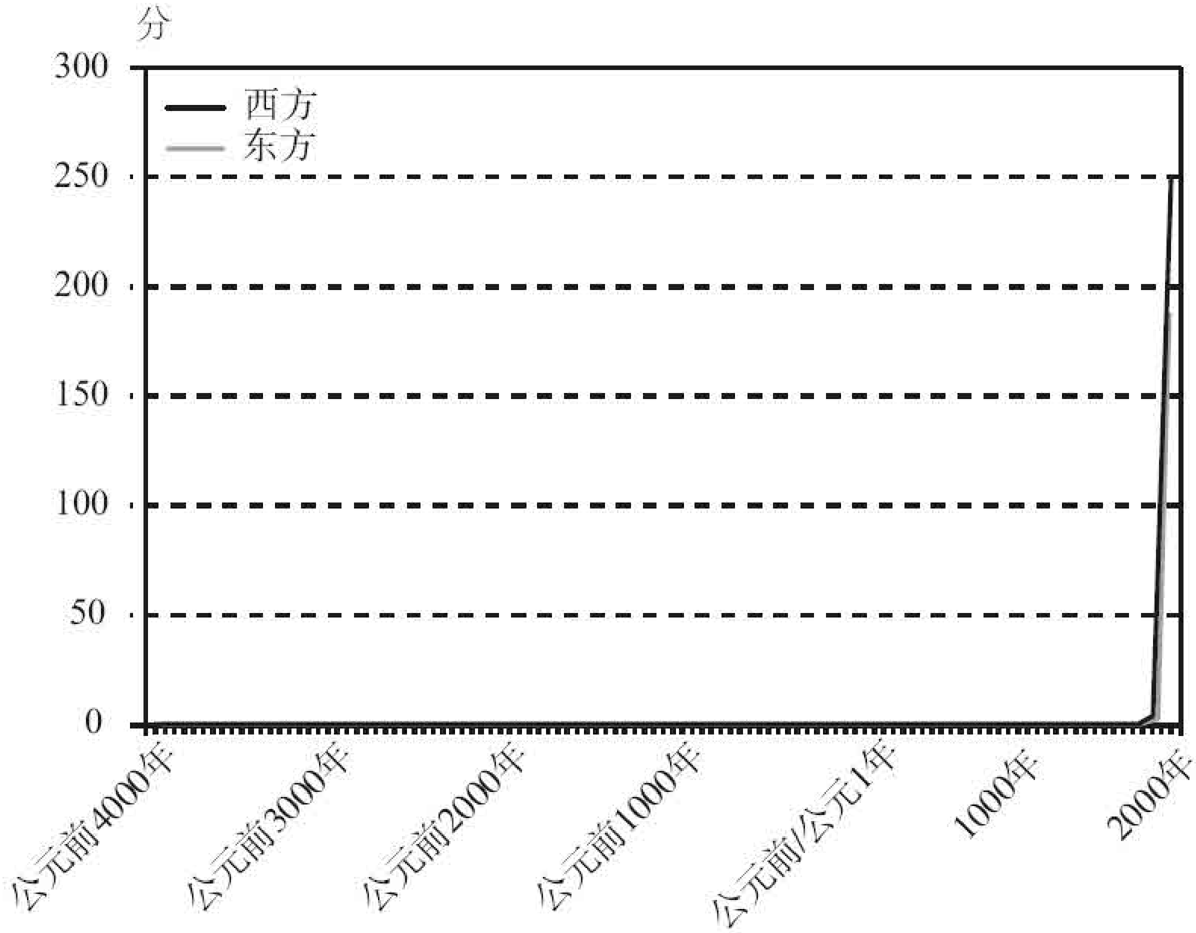
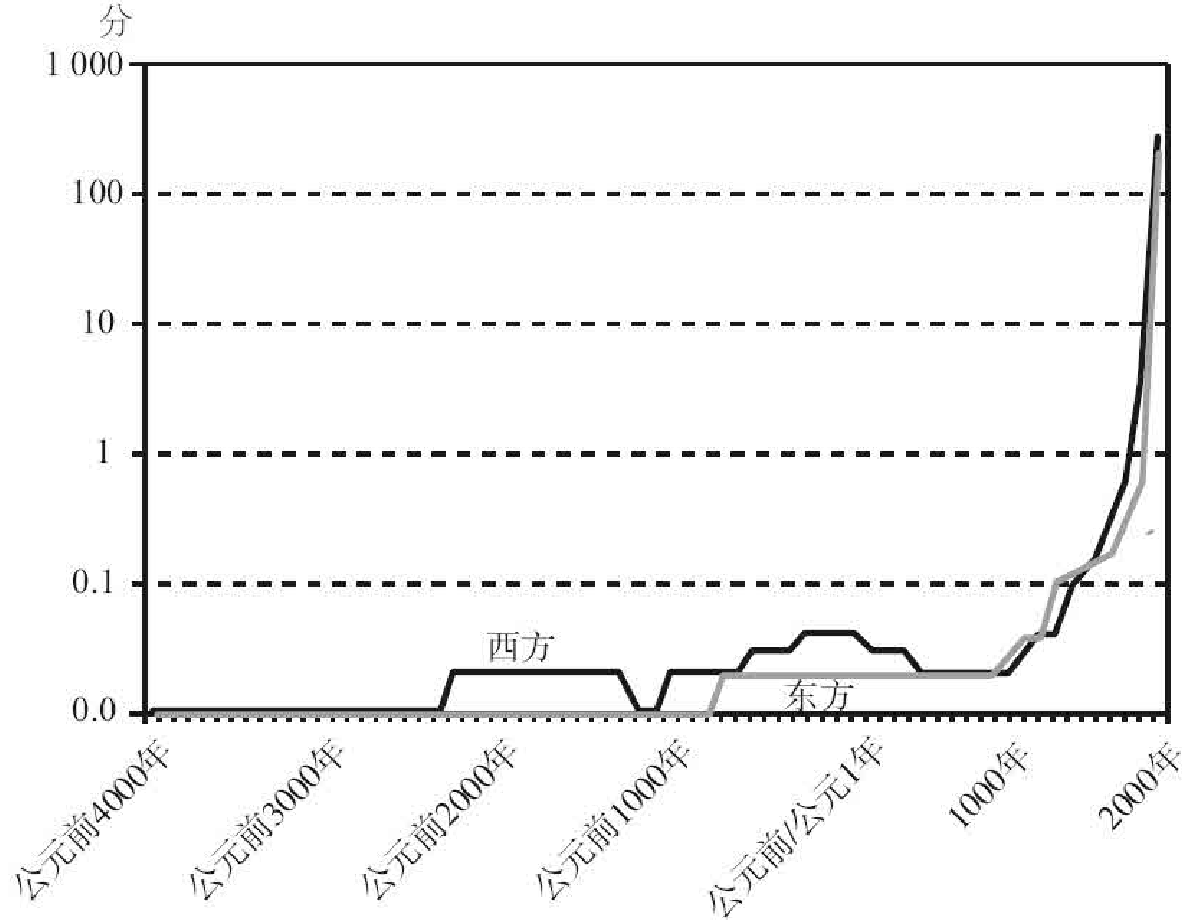


图6.4 公元前4000~公元2000年东方和西方的信息技术(以对数–线性标尺标绘，东方1400~1900年，西方1500~1800年因印刷术而分数翻番）



这种计算方法依赖于一个更进一步的关键假设：采用可见符号来记录概念是至关重要的。人类交谈和计算几万年后，才开始书写或使用数字符号，我们通过传统、仪式和艺术储存和交流了浩如烟海的信息。然而，很显然的是，所有纯粹口头系统的信息技术，在社会发展指数中自然都是零分。

我这样做有三个理由。

第一，出于生物学的考虑：人脑都是一样的，尽管先前曾提到有人声称文化之间有极端变量，但至今尚无令人信服的证据表明不同口头文化的人群在头脑中处理和存储信息，或者通过说话交流信息的能力有什么重大不同。如果这种认识是正确的，出于比较的目的，文字

产生以前的信息技术系统实际上都要归零。只有更成熟的文字和计算技术开发出来后，可以衡量的差异才开始出现。

第二，出于实际的考虑：即使上一段所提到的假设是错误的，我也不知道有什么办法来衡量和比较古代没有文字的不同文化的信息技术系统。假如在符号系统最早的证据在两个地区出现（西方为大约公元前9300年，东方为公元前7000年）之前的时代，东方的口头文化在处理、储存和交流信息方面是否强于西方的口头文化，我们无从得知。

第三，出于经验的考虑：使用可见符号来记录述谓性概念和数学概念的革命性的影响已经非常确定了。经常给那些强调直观记录的有效性的人贴上“进化论者”标签的批评家们，提出了大量理由要求人们当心极端的主张，在解释文字的影响时要持灵活态度；但是在经过了半个世纪的争论后，从纯粹口头转变为口头和文字的信息技术的各种组合，究竟是加强了个人能力，产生了等级层次，还是同时促成了两者，似乎仍然不清楚，这也标志着在增强人的存储、使用和传播信息的能力方面迈出了一大步。在西方，这方面的证据得到了极其详细的研究，最早的符号可能是用来记账的，从中也逐渐产生了口头形式。在东方，证据不那么清楚，但同样的模式应当也适用。

表6.1 西方信息技术得分

年代	充分（0.5分）	中等（0.25分）	初步（0.15分）	男性分数（分）	女性（%男性）	读写分数（分）	乘数	总分（分）
公元2000年	100 (50)	0	0	50	100% = 50	100	2.5	250
公元1900年	40 (20)	50 (12.5)	7 (1.05)	33.6	90% = 30.2	63.8	0.05	3.19
公元1800年	20 (10)	25 (6.25)	20 (3)	19.3	50% = 9.65	28.95	0.01	0.29
公元1700年	10 (5)	15 (3.75)	25 (3.75)	12.5	10% = 1.25	13.75	0.01	0.14

公元1600年	5 (2.5)	10 (2.5)	10 (1.5)	6.5	2% = 0.13	6.63	0.01	0.071
公元1500年	4 (2)	8 (2)	6 (0.9)	4.9	2% = 0.10	5.0	0.01	0.05
公元1400年	3 (1.5)	6 (1.5)	4 (0.6)	3.6	1% = 0.04	3.64	0.01	0.04
公元1300年	3 (1.5)	6 (1.5)	4 (0.6)	3.6	1% = 0.04	3.64	0.01	0.04
公元1200年	3 (1.5)	6 (1.5)	4 (0.6)	3.6	1% = 0.04	3.64	0.01	0.04
公元1100年	2 (1)	4 (1)	2 (0.3)	2.3	1% = 0.02	2.32	0.01	0.02
公元1000年	2 (1)	4 (1)	2 (0.3)	2.3	1% = 0.02	2.32	0.01	0.02
公元600~900年	2 (1)	2 (0.5)	1 (0.15)	1.65	1% = 0.02	1.67	0.01	0.02
公元300~500年	3 (1.5)	4 (1)	3 (0.45)	2.95	1% = 0.03	2.98	0.01	0.03
公元前100~公元200年	4 (2)	6 (1.5)	5 (0.75)	4.25	1% = 0.04	4.29	0.01	0.04
公元前500~前200年	2 (1)	3 (0.75)	2 (0.3)	2.05	1% = 0.02	2.07	0.01	0.02
公元前900~前600年	1 (1)	2 (0.5)	1 (0.15)	1.65	1% = 0.02	1.67	0.01	0.02
公元前1100~前1000年	1 (1)	1 (0.25)	1 (0.15)	1.4	1% = 0.01	1.41	0.01	0.01
公元前2200~前1200年	1 (1)	2 (0.5)	1 (0.15)	1.65	1% = 0.02	1.67	0.01	0.02
公元前2700~前2300年	1 (1)	1 (0.25)	1 (0.25)	1.4	1% = 0.01	1.41	0.01	0.01
公元前3300~前2800年	0 (1)	1 (0.25)	2 (0.3)	0.55	1% = 0.01	0.56	0.01	0.01
公元前6000~前3400年	0	0	1 (0.15)	0.15	1% = 0	0.15	0.01	0
公元前9000~前6100年	0	0	0	0	0	0	0.01	0

公元前9300~前9000年	0	0	1 (0.15)	0.15	1% = 0	0.15	0.01	0

表6.2 东方信息技术得分

年代	充分（0.5分）	中等（0.25分）	初步（0.15分）	男性分数	女性（%男性）	读写分数	乘数	总分（分）
公元2000年	100 (50)	0	0	50	100% = 50	100	1.89	189.00
公元1900年	15(7.5)	60 (15)	10 (1.5)	24	25% = 6	30	0.01	0.30
公元1800年	5 (2.5)	35 (8.75)	10 (1.5)	12.75	5% = 0.64	13.39	0.01	0.13
公元1700年	5 (2.5)	20 (5)	10 (1.5)	9	2% = 0.18	9.18	0.01	0.09
公元1600年	4 (2)	15 (3.75)	10 (1.5)	7.25	2% = 0.15	7.4	0.01	0.07
公元1500年	3 (1.5)	10 (2.5)	10 (1.5)	5.5	2% = 0.11	5.61	0.01	0.06
公元1400年	3 (1.5)	10 (2.5)	10 (1.5)	5.5	2% = 0.11	5.61	0.01	0.06
公元1300年	3 (1.5)	5 (1.25)	5 (0.75)	3.5	1% = 0.04	3.51	0.01	0.04
公元1200年	3 (1.5)	5 (1.25)	5 (0.75)	3.5	1% = 0.04	3.51	0.01	0.04
公元1100年	2 (1)	2 (0.5)	3 (0.45)	1.95	1% = 0.02	1.97	0.01	0.02
公元前600~公元1000年	2 (1)	2 (0.5)	2 (0.3)	1.8	1% = 0.02	1.82	0.01	0.02
公元前1000~前700年	2 (1)	1 (0.25)	1 (0.15)	1.4	1% = 0.01	1.14	0.01	0.01
公元前1300~前1100年	1 (0.5)	1 (0.25)	1 (0.15)	0.9	1% = 0.01	0.91	0.01	0.01
公元前7000~前1400年	0	0	1 (0.15)	1.15	1% = 0	0.15	0.01	0

西方识字率

当我们沿时间长河回溯时，证据的性质变化极大，但我们仍然能拼出一幅非常粗略的画面。20世纪60年代中期到80年代中期，历史学家对1600~1900年欧洲的识字率做了些先驱性的探索工作，讨论了不同时间男性和女性识字的不同水平。也有一些学者对美国开展了少量的同类工作。

自20世纪80年代中期起，这类统计方法遭到了批评，历史学家逐渐放弃了有利于文化史研究图书和读者群体的量化方法。在重构现代早期识字率时，所运用的方法问题无疑是重要的，但研究方法的改变，似乎更多的是受背离定量分析的更广泛的史学趋势的影响，而不是因为有确凿的证据表明20世纪60~80年代的研究结果有什么严重的错误。

从专家研究中浮现出的整体画面是，自1600年以来整个欧洲和北美洲识字率的地区变化和广泛趋势相结合，各级水平的识字率都有所提高，并且男女之间的差距也在缩小。在我的指数上，将奇波拉（Cipolla）、斯通（Stone）及其他人提出的数字转换成分数后，表现出自1600~1800年大致每世纪翻了一番，（社会发展分数）从1600年的0.07分增长至1800年的0.29分，然后又飙升至1900年的3.19分。

公元1600年之前，证据情况不大好。中世纪史专家大张旗鼓地研究了欧洲关于读写情况的文献，但对计算情况关注不够。在伊斯兰核心地带，情况却正好相反，关于读写情况的文字资料极少，但自然科学和数学得到了更多的关注。也很少有人专注于研究中世纪的伊斯兰教育情况和更广大的伊斯兰人口的读写和计算能力。

自历史学家们有时称之为“12世纪复兴”的时代起，西欧男性的读写能力和计算能力开始缓慢增强，而在公元1100年前，他们的水平非

常低，对此意见似乎较一致。会读写和计算的女性的数量，可能是直到1500年后才开始稳步增长的。

研究伊斯兰教育的学者基本上无人愿意冒险去做任何量化估计，但看来公元1100年前，尽管顶级的伊斯兰学者计算能力比基督教世界顶级学者要强，读写能力至少也不相上下，但读写却局限于极少数人。我们也许可以说，中世纪伊斯兰的读写工作只是抄写员和神职人员的绝技，而在基督教的欧洲，读写却为更广大的工匠阶层普遍掌握（尽管能读的读物通常都与《圣经》相关）。伊斯兰世界没有出现16世纪的欧洲男性踊跃阅读宗教经典并带动识字女性增多的兴旺景象。

1100年时，仅仅具备最基本的阅读水平的西方人可能还不到10%，算得上有完全读写能力的人甚至更少（可能仅有2%）。能读会写的女性数量尤其难以知晓，但有文化的女性与有文化的男性的比率恐怕要达到1:100，实在少到几乎无法对分数产生影响。我估计大约1100年时西方的社会发展分数只有0.02分，到1500年时以缓慢的增量达到了0.05分，然后增长速度才得以加快。

读写和计算在古典时代比在中世纪，似乎普及更广、更深，尤其是在民主时期的雅典（公元前508~前322年）和大约公元前200~公元200年的意大利。在此方面，威廉·哈里斯（William Harris）做出过非常过硬的量化估计，我基本上将采用他的数字。

新近对古希腊–罗马读写水平的研究，像中世纪史专家和人类学家们一样，强调读写是一种非常复杂的现象，不可能由一个单一的分数来反映，但哈里斯在计算比率时，已经将读写的各种各样的形式考虑在内了。

结合这些批评，我估计西方核心地带信息技术的社会发展分数在公元前100~公元200年这一时期达到的巅峰为大约0.04分。公元200年后开始下跌，我为公元300~500年估计的分数为0.03分，此后，由于缺

乏任何更清楚的证据，也因为所涉及的数字实在太小，我给直到公元1100年复兴之前所打的分数，都停留在0.02分这个静止的水平上。

再回望公元前100年之前，公元前400~前200年，在爱琴海岸和东地中海的核心地带，信息技术得0.03分，与公元前第一个千年早期的0.02分比，有所增长。分数这样微小，精确性和细微差别都谈不上了；我给公元前2200年（阿卡德王朝和乌尔第三王朝等官僚国家兴起）到公元前500年（希腊的民主国家开始扩张）的信息技术打的分数基本上是持平的，表示历史学家经常称之为“抄写识字文化”和“工匠识字文化”（我还要加上“识数文化”）的结合。

说到抄写/工匠识字文化，我指的是有一个人数极少的受过教育的精英阶层（大概只占男性人口的1%）完全掌握了文学经典，一个人数稍多一些（大概占男性人口的2%）的官僚精英阶层掌握了记录技术，还有一个较小的工匠群体（也许是1%~2%）能够读和写自己的名字，并且能进行他们的专业所需要的计算。这种抄写/手艺信息技术所得的社会发展分数为0.02分，除了在公元前1200~前1000年的崩溃时期出现了一次中断，各种文字证据都骤然萎缩。在希腊，文字可能完全无用了，在整个东地中海一带，只有非常少的文字资料留存了下来。对那个“黑暗时代”，我打0.01分。

抄写识数文化和识字文化最早的确凿证据，大约于公元前3300年出现在美索不达米亚南部，我给这段时期打0.01分。信息技术在接下来的1 000年中成熟度和应用范围均有所增长，但鉴于社会发展指数中可用的最小增量是0.01分，图6.2中对这段时期只能用平直线标示，直到大约公元前2250年才向上跳跃。早在公元前9000年也有人类符号活动的细微迹象，也有一些学者称之为书写或算数，但这些痕迹实在太稀少，我仍给它们打零分。

东方识字率

在我能看懂的语言中，对东方识字文化和识数文化的定量分析比对西方的要少得多，图6.2中的平直分数线和表6.2的简略便反映了这一点。我给东方打的分数必然将复杂得多的模式过分简化了，使之像西方的分数图中标绘的一样，充满了起伏。

公元2000年，我采用HDI，将日本核心地带的东方识字率视为与西方核心地带的相同，但正如我在“信息技术记分牌”一节中所解释的，我使用乘数1.89，而不是西方的2.5，来表示2000年时日本电子媒介的可用性比美国要狭窄。

1900年，日本政府开始为提高识字率付出艰辛努力。虽然标准与西方核心地带相比算是低的，但比前现代文化要高得多，可能有85%的男孩和25%的女孩至少掌握了一些技能。就日本达到的水平，存在一些争议，但由于日本的信息技术即使在1900年，仍然在很大程度上是电之前的技术，所以在这个时期，东西方在社会发展分数上的差距是巨大的。

我计算的东方分数是0.3分（30 ITP ×0.01，表示仍处于电之前时期），而西方的分数是3.19分。1900年左右，中国人的识字和识数水平比日本还要低，因为受过教育的精英阶层对于大众教育存在矛盾心态。中国人的文化程度按照前现代的标准算是非常高了，1900年时可能至少有50%的男孩达到了基本的识字标准，但迈向大众教育的步伐仍然相当迟疑。直到1949年中华人民共和国成立后，大众教育才真正起飞。

1900年之前，清朝时期的中国基础教育和工匠识字文化在稳步扩大。1700年左右，可能只有5%的男人可以说能够比较流畅地阅读，35%的男孩粗通文墨，但到1800年时，中国北方将近一半的男孩都达到了粗通文墨的水平。

女性读写和识数受到很大限制。18~19世纪西方人的识字和识数率比东方人高（尤其是女性），但绝对数字仍然很小，社会发展分数的实际差别（按我的计算，西方1700年为0.14分，1800年翻番为0.29分，东方1700年为0.09分，1800年时大致增长了50%，为0.13分）并不大。

明朝时期的分数似乎更低，尽管可能比1600年开始信息繁荣之前的西方要高。在欧亚大陆两端，受过良好教育的精英阶层可能没多大区别，但中国似乎有一个相当大的中等识字识数文化程度的群体（男性占压倒多数）。实际分数必然是凭印象打的（我给1500年打0.06分，给1600年打0.07分，西方同期分别为0.05分和0.07分），然而由于17世纪前的分数如此之低，除非误差非常大才会对社会发展指数产生严重影响。当时日本的水平可能与中国非常接近。

回溯到更早的时期，打分的精确性无疑会更差。精英教育在唐朝和宋朝得到了很大改善，10~12世纪，书籍和财务记录始终处于爆炸状态，使我感到信息技术的应用大致堪与西方的罗马帝国时期（0.04分）媲美。打0.03分或0.05分似乎也同样合理，但低至0.02分（相当于我给公元600~900年的西方打的分数）或高达0.06分（相当于16世纪西方的分数）恐怕都是不可能的。分数从1000年的0.02分迅速上升到1400年的0.06分。

由于没有更好的理由做其他选择，我索性假设公元前600~公元1000年这段漫长的时期都是同一个分数0.02分。在这16个世纪，识字率和识数率无疑都是波动的。定性证据表明，两者可能在公元前600~公元100年都上升了，在公元100~400年都下降了，然后在公元400年后又再次上升了。越来越多的汉朝文献正在被发现，特别是在干旱的西北和湿润的南方，气候条件利于竹简留存。然而，目前看来，中国人的识字率和识数率似乎仍然低于罗马的水平。汉朝灭亡之后信息技术的衰退似乎也不像西方后罗马时代的衰退那样严重。中国在0.02分上

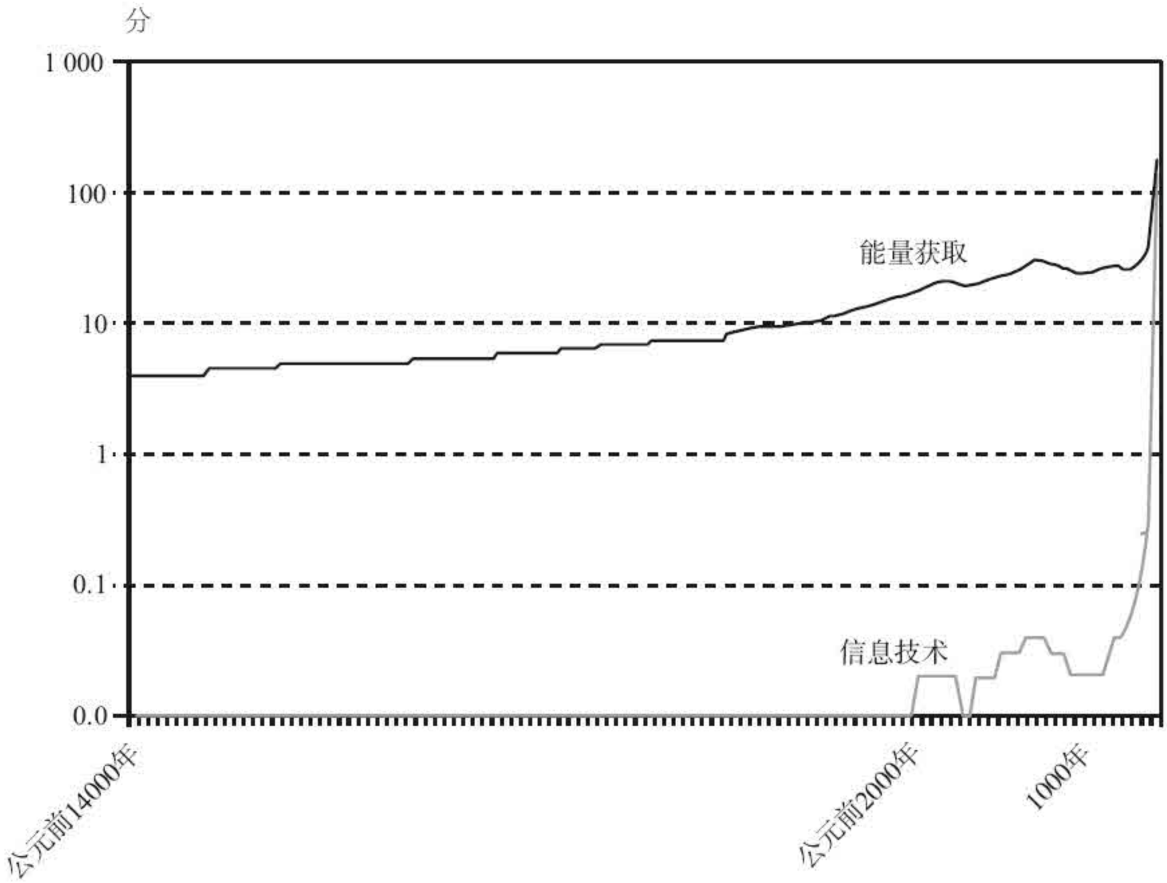
下的变量从历史上看一定很重要，但在社会发展指数上，恐怕就小得可以忽略不计了。

中国符号表示法最早的证据来自大约公元前6250年的贾湖，也有充分证据显示在接下来的5 000年这一实践存在着连续性。然而，直到大约公元前1300年，中国使用的文字符号和数学符号似乎才能够与公元前3000年左右美索不达米亚的符号相媲美，能够得到0.01分。在接下来的1 000年里，证据显示符号系统的应用不断扩张，从甲骨文到青铜器皿上的铭文，再到大规模地使用墨在竹简和丝绸上书写。然而，分数实在是太低了，这些改进在社会发展指数上，只能从0.01分跳到我给大约公元前600年打的0.02分。

信息技术的反馈回路

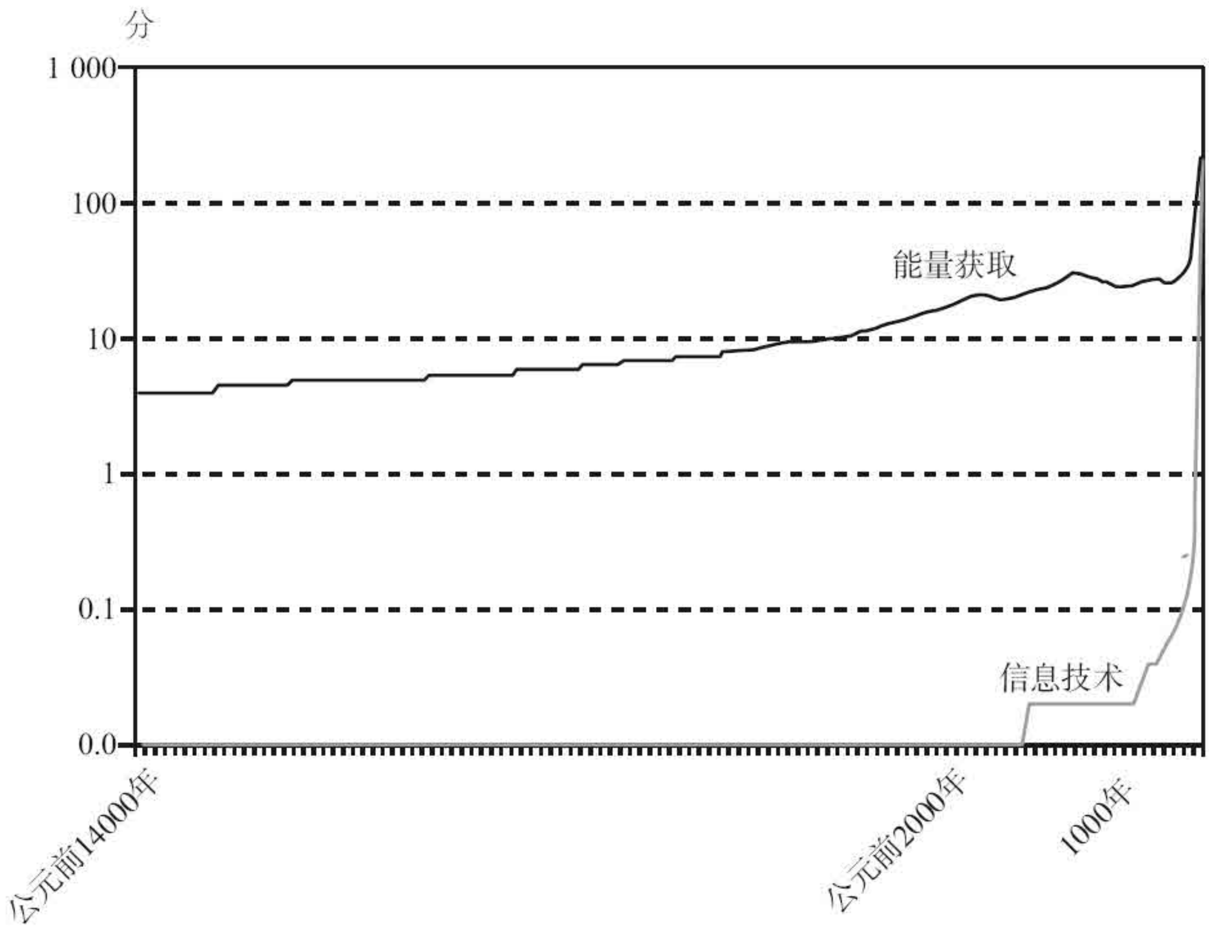
图6.5和图6.6非常清楚地表明，无论在东方还是在西方，信息技术都对过去几个世纪广泛的变化格外敏感。

图6.5 公元前14000~公元2000年西方能量获取水平与信息技术在对数－线性标尺上的对比（以社会发展分数衡量）



信息技术和能量获取被卷入了一个反馈回路。如果人们的读写能力和计算能力没有发展到一定的水平，18世纪晚期最初的英国工业革命就不可能发生。而19世纪晚期使化学全面工业化的“第二次工业革命”，更是倚重于信息技术。在我们自己这个时代，20世纪末21世纪初生产力的爆炸式发展，与全新形式的信息技术的彻底腾飞，关系更是极其密切。

图6.6 公元前14000~公元2000年东方能量获取水平与信息技术在对数－线性标尺上的对比（以社会发展分数衡量）



新近的信息爆炸意味着1700年之前的所有信息技术得分都必须极低。信息技术是社会发展指数4个特性中最难衡量的，但由于前现代时期的分数实在太低，即使有误差，似乎也不大可能对总体的社会发展分数造成明显差异。

The Measure of Civilization

How Social Development Decides the Fate of Nations

第七章

让历史成为指南：度量文明的意义

在本书中，我介绍了支持一种分析工具——社会发展指数——的证据和方法。因此，以一个关于这种工具能够做什么、不能做什么的更开放的讨论，而不是一系列结论来结束本书，似乎是明智的。

我先用两节来讨论一下指数可能出现的问题。首先，我对误差和讹误做些评论。新进化论者的指数最大的缺陷之一是，由于指数并非真正为回答某个具体问题而设立，其设计者很难准确地说出它们为什么会出现讹误。误差项取决于所问的问题，就为什么西方主宰世界这个问题，我们对于能够容忍多大的误差才能保证指数不误导，可以做到相当准确。

在接下来的一节中，我将转向展示数据的问题。在展示统计信息时，根本没有中立的方式；各种方式都会强调指数的一个或多个维度而忽视其他维度。我成系统地选择了最简单的方式，在一切可能的情况下均优先选用线性–线性图，但其他方式也有其优点。

然后我将回到在第一章中曾考虑的问题上，探讨此处谈论的指数怎么会有助于统一历史上的进化理论。我对有可能形成一个这样的理论，而且社会发展指数可以成为这个理论的一个重要组成部分，持乐观态度。最后，我将探讨这样一个理论能否真的使历史成为未来的指南，就像斯宾塞150多年前所声称的那样。

误差和讹误

第三章至第六章讨论过的几乎所有细节都可能受到质疑，证据几乎都可以进行多种解释。有不止一种方式来定义能量获取、社会组织、战争能力和信息技术。我也能以其他方式计算分数。在产生社会发展指数的每个阶段，都会牵扯出长长的一连串争论和推断。

因此，另一位研究者完全可能得出一套完全不同的社会发展分数。实际上，任何其他研究者是极其不可能得出与我完全一致的一套分数的。就此而言，如果我再重新计算一遍社会发展分数，我本人也很可能得出不同的数字。

所以，问指数是否正确这个问题，是没有用的。从来没有指数是“正确”的，无论我们是从严格的意义上理解为表7.1和表7.2中的所有530个数字都完全符合实际，还是从宽泛的意义上理解为所有专家都将同意这些数字。我计算的分数肯定是错误的，唯一有用的问题是它们错得有多厉害。它们是错到了连图2.5中描绘的历史基本轮廓都属误导，以致整个《西方将主宰多久》这本书都有致命缺陷呢？还是误差实际上微不足道?

想确知这些问题的答案，唯一办法是让其他考古学家和历史学家来研究我在第三章到第六章中展示的证据，并检验我的论点。不过，在此，我至少能肯定地说出指数能容忍多大的分数错误，以及什么情况能促使我改变主张。如果我计算的分数基本上和其他分析家计算的数字相差在10%以内，那么我试图解释的模式的基本轮廓就仍然可以维持现状。如果我的数字基本上和其他人差15%，那么我也许要将发展曲线的形状改变得足以修正我的观点，这取决于细节。如果相差在20%以上，那么显然我的观点需要修正了。

表7.1 公元前14000~公元2000年西方社会发展分数

（单位：分）

年代	能量获取	社会组织	战争能力	信息技术	总分
公元前14000年	4.36	0.00	0.00	0.00	4.36
公元前13000年	4.36	0.00	0.00	0.00	4.36
公元前12000年	4.90	0.00	0.00	0.00	4.90
公元前11000年	5.45	0.00	0.00	0.00	5.45

公元前10000年	5.45	0.00	0.00	0.00	5.45
公元前9000年	5.99	0.00	0.00	0.00	5.99
公元前8000年	6.54	0.00	0.00	0.00	6.54
公元前7000年	7.08	0.01	0.00	0.00	7.09
公元前6000年	7.63	0.03	0.00	0.00	7.66
公元前5000年	8.72	0.04	0.00	0.00	8.76
公元前4000年	10.90	0.05	0.00	0.00	10.95
公元前3500年	11.99	0.09	0.00	0.00	12.98
公元前3000年	13.08	0.42	0.01	0.01	13.52
公元前2500年	15.26	0.47	0.01	0.01	16.29
公元前2250年	17.44	0.33	0.01	0.01	17.79
公元前2000年	18.52	0.56	0.01	0.02	19.11
公元前1750年	20.65	0.61	0.02	0.02	21.30
公元前1500年	22.34	0.70	0.03	0.02	23.09
公元前1400年	22.88	0.75	0.03	0.02	23.68
公元前1300年	23.43	0.75	0.03	0.02	24.23
公元前1200年	22.88	0.75	0.02	0.04	23.69
公元前1100年	22.34	0.47	0.03	0.01	22.85
公元前1000年	21.79	0.47	0.03	0.01	22.30
公元前900年	22.34	0.47	0.04	0.02	22.87
公元前800年	22.88	0.70	0.05	0.02	23.65
公元前700年	23.43	0.94	0.07	0.02	24.45
公元前600年	23.97	1.17	0.07	0.02	25.23
公元前500年	25.06	1.40	0.08	0.03	26.56
公元前400年	26.15	1.40	0.09	0.03	27.67
公元前300年	28.33	1.40	0.09	0.03	29.85
公元前200年	29.42	2.81	0.10	0.03	32.36
公元前100年	31.06	3.75	0.11	0.04	35.50
公元前/公元1年	33.78	9.36	0.12	0.04	43.30

公元100年	33.78	9.36	0.12	0.04	43.30
公元200年	32.69	9.36	0.11	0.04	42.20
公元300年	31.60	7.49	0.10	0.03	39.22
公元400年	31.60	7.49	0.09	0.03	38.67
公元500年	30.51	4.23	0.07	0.03	34.84
公元600年	28.33	1.41	0.04	0.02	29.80
公元700年	27.24	1.17	0.04	0.02	28.47
公元800年	27.24	1.64	0.04	0.02	28.94
公元900年	27.24	1.64	0.05	0.02	28.95
公元1000年	28.33	1.87	0.06	0.02	30.28
公元1100年	28.33	2.34	0.07	0.02	30.76
公元1200年	28.88	2.34	0.08	0.03	31.33
公元1300年	29.42	3.75	0.09	0.04	33.31
公元1400年	28.33	1.17	0.11	0.04	29.65
公元1500年	29.42	3.75	0.13	0.05	33.35
公元1600年	31.06	3.75	0.18	0.07	35.60
公元1700年	34.87	5.62	0.35	0.14	40.98
公元1800年	41.41	8.43	0.50	0.29	50.63
公元1900年	100.25	61.80	5.00	3.19	170.24
公元2000年	250.00	156.37	250.00	250.00	906.37

按照图7.1中以对数–线性标尺展示出的指数，西方的社会发展水平自公元前14000年后领先于东方。东方慢慢追赶了上来，尤其是在公元前2000年后，在公元前第一个千年的大部分时间里，西方的领先优势都很小。公元前100年左右，西方再次将领先优势扩大。但到公元541年时，东方首次升到了西方之上。东方的分数保持领先直到1773年。自冰期结束后92.5%的时间里，西方的发展水平均高于东方。

图7.1 公元前14000~公元2000年东方和西方的社会发展分数（以对数－线性标尺标绘）

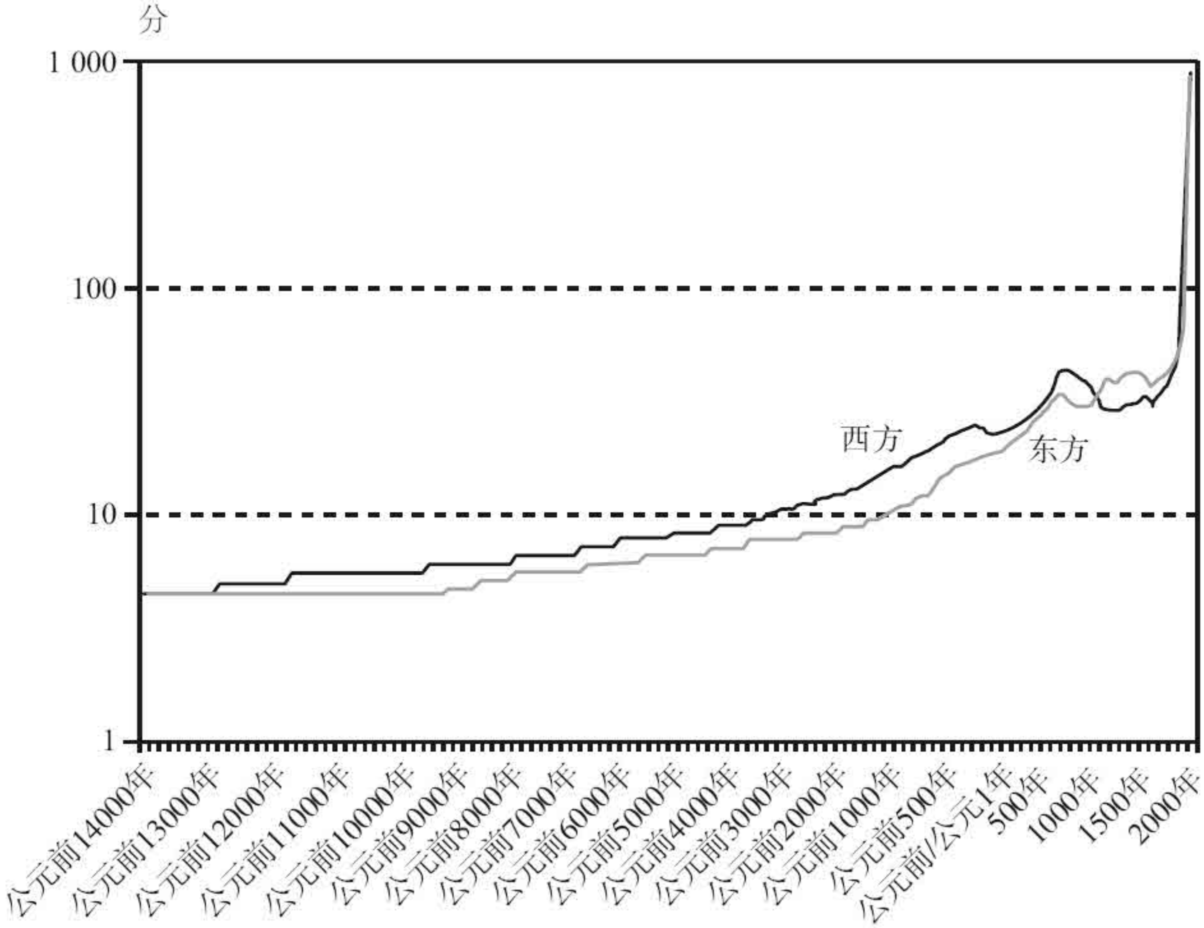


表7.2 公元前14000~公元2000年东方社会发展分数

（单位：分）

年代	能量获取	社会组织	战争能力	信息技术	总分
公元前14000年	4.36	0.00	0.00	0.00	4.36
公元前13000年	4.36	0.00	0.00	0.00	4.36
公元前12000年	4.36	0.00	0.00	0.00	4.36
公元前11000年	4.36	0.00	0.00	0.00	4.36
公元前10000年	4.36	0.00	0.00	0.00	4.36
公元前9000年	4.90	0.00	0.00	0.00	4.90
公元前8000年	5.45	0.00	0.00	0.00	5.45
公元前7000年	5.99	0.00	0.00	0.00	5.99

公元前6000年	6.54	0.00	0.00	0.00	6.54
公元前5000年	7.08	0.00	0.00	0.00	7.08
公元前4000年	7.63	0.00	0.00	0.00	7.63
公元前3500年	8.17	0.02	0.00	0.00	8.19
公元前3000年	8.72	0.05	0.00	0.00	8.77
公元前2500年	10.35	0.09	0.00	0.00	10.44
公元前2250年	11.44	0.13	0.00	0.00	11.57
公元前2000年	11.99	0.10	0.00	0.00	12.09
公元前1750年	14.17	0.22	0.00	0.00	14.39
公元前1500年	16.35	0.33	0.01	0.00	16.69
公元前1400年	16.89	0.33	0.01	0.00	17.23
公元前1300年	17.44	0.33	0.01	0.01	17.79
公元前1200年	17.44	0.47	0.02	0.01	17.94
公元前1100年	17.98	0.47	0.02	0.01	18.48
公元前1000年	18.52	0.33	0.03	0.01	18.89
公元前900年	19.07	0.37	0.03	0.01	19.48
公元前800年	19.61	0.42	0.02	0.01	20.06
公元前700年	20.16	0.51	0.02	0.01	20.70
公元前600年	21.79	0.61	0.03	0.02	22.45
公元前500年	22.88	0.75	0.04	0.02	23.69
公元前400年	23.97	0.94	0.05	0.02	24.98
公元前300年	24.52	1.17	0.06	0.02	26.87
公元前200年	26.15	2.81	0.07	0.02	29.05
公元前100年	27.79	3.45	0.08	0.02	31.64
公元前/公元1年	29.42	4.68	0.08	0.02	34.20
公元100年	29.42	3.93	0.08	0.02	33.44
公元200年	28.33	1.12	0.07	0.02	29.54
公元300年	28.33	1.31	0.07	0.02	29.73
公元400年	28.33	1.87	0.07	0.02	29.99

公元500年	28.33	1.87	0.08	0.02	30.30
公元600年	29.42	5.63	0.09	0.02	35.16
公元700年	29 .42	9.36	0.11	0.02	38.91
公元800年	30.51	9.36	0.07	0.02	39.96
公元900年	31.06	7.00	0.07	0.02	38.69
公元1000年	32.15	9.36	0.08	0.02	41.61
公元1100年	32.69	9.36	0.09	0.02	42.17
公元1200年	33.23	9.36	0.09	0.03	42.71
公元1300年	32.69	7.50	0.11	0.04	40.34
公元1400年	31.06	4.68	0.12	0.05	35.91
公元1500年	32.69	6.35	0.10	0.06	39.20
公元1600年	33.78	6.55	0.12	0.07	40.52
公元1700年	35.96	6.09	0.15	0.09	45.29
公元1800年	39.23	10.30	0.12	0.13	49.78
公元1900年	53.40	16.39	1.00	0.30	71.09
公元2000年	113.33	250.00	12.50	189.00	564.83

图7.2在对数－线性标尺上展示了假如我将西方的发展分数一直低估了10%，而将东方的发展分数高估了10%，那么东方和西方正确的走势将是怎样的（即在图中将我计算的所有西方分数提高10%，将所有东方的分数降低10%）。图7.3则展示了假如我犯了相反的错误，将东方发展分数低估了10%而将西方发展分数高估了10%，将会是怎样的。

首先要注意的一点是图7.2和图7.3的可信度扭曲了多少。图7.2分别将西方和东方的分数提高和降低了10%，这就要求我们承认在公元1400年，当郑和正准备扬帆印度洋时，西方比东方要发达；这也意味着当公元前218年，汉尼拔驱赶着他的大象进攻罗马时，西方的发展水平已经比东方的郑和时代要高了。好像这些结论还不够荒唐似的，该

图还告诉我们，当公元前44年恺撒遇刺时，西方已经比公元1793年乾隆皇帝拒绝马戛尔尼勋爵的通商请求时的中国要发达了。所有这些结论没有一个能与众多历史证据相符。

图7.2 公元前14000~公元2000年东方和西方的社会发展分数（以对数–线性标尺标绘，将所有西方的分数提高10%，将所有东方的分数降低10%）

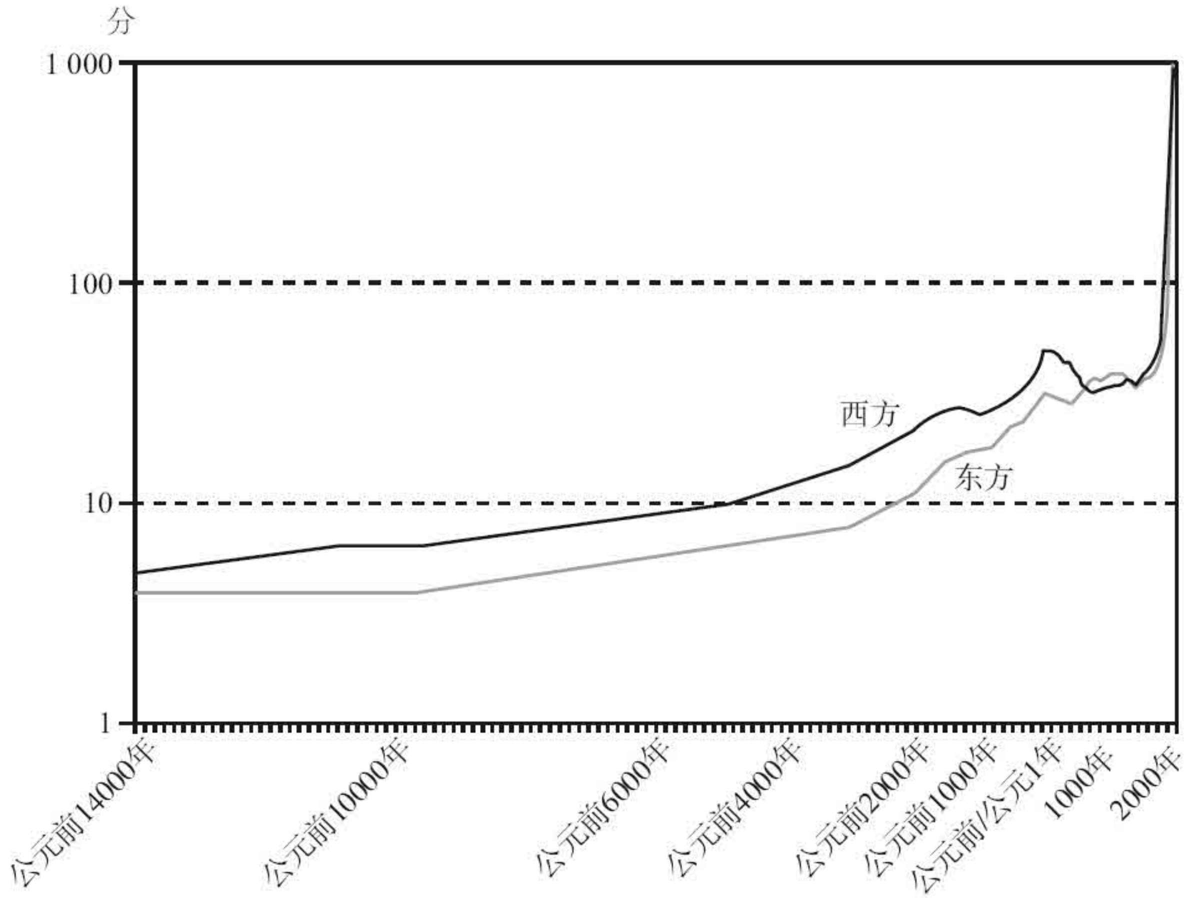


图7.3 公元前14000~公元2000年东方和西方的社会发展分数（以对数－线性标尺标绘，将所有西方的分数降低10%，将所有东方的分数提高10%）

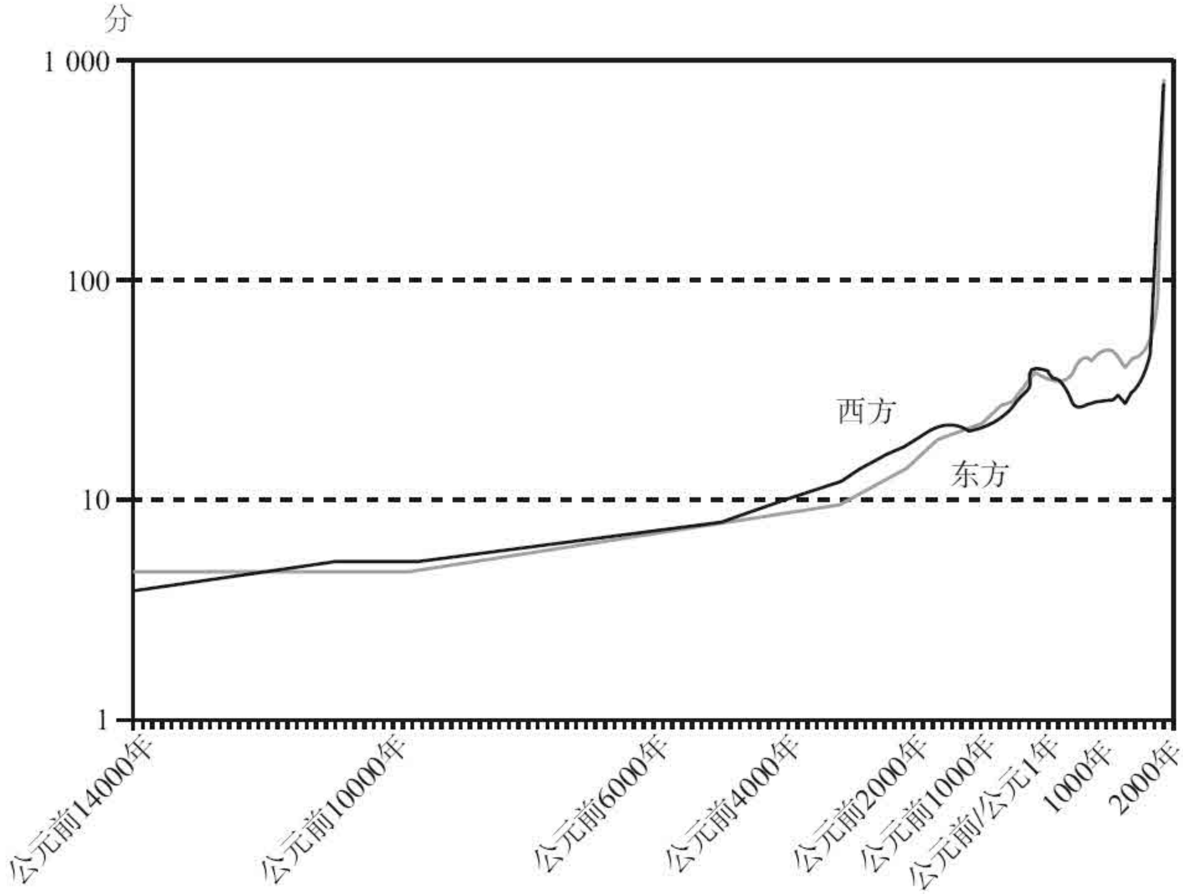


图7.3将东方的分数提高了10%而将西方的分数降低了10%，甚至更为荒唐。例如，700年时，大马士革的哈里发统治着一个从今葡萄牙一直延伸到巴基斯坦的大帝国，该图给当时西方打的发展分数却比孔子时代的东方还要低，这恐怕根本不可能；1800年，工业革命已经展开，英法两大帝国攫取了全球大片土地，然而该图给西方打的分数，却比东方在1000~1200年的宋朝时期还要低，似乎更令人难以置信。

然而，即使历史学家们能够忍受这样奇怪的结论，图7.2和图7.3所描绘的历史的形状与图7.1所描绘的，差异并没有大到要改变基本模式的地步。短期偶然论仍然不足以成立，因为即使在图7.3中，西方的分数在冰期结束后的大多数时期仍然高于东方（只不过这个“大多数”意味的是56%的时间，而不是92.5%）。长期注定论也不足以成立，因为即使在图7.2中，东方也领先了7个世纪。我计算的分数所产生的模式

——西方在过去15 000年的大部分时间里处于领先，但中间插入了一个1 200年的“东方时代”——仍然是不变的。

要改变基本模式，我估计必须是在误差达到了20%以上的情况下。图7.4展示了假如我对西方的发展分数一直低估了20%，而对东方的分数高估了20%，那么历史将是什么样子。图7.5则展示了假如我将东方的发展分数低估了20%而将西方的分数高估了20%的结果。

这回模式大不相同了。在图7.4中，西方的分数总是高于东方，使得长期注定论貌似非常合理，也否定了我在《西方将主宰多久》一书中通篇的主张：社会发展改变了地理的意义。图7.5则相反，实际上颠倒了我的实际指数得出的结论，使东方在冰期以后90%的时间内处于领先地位。

图7.4 公元前14000~公元2000年东方和西方的社会发展分数（以对数–线性标尺标绘，将所有西方的分数提高20%，将所有东方的分数降低20%）

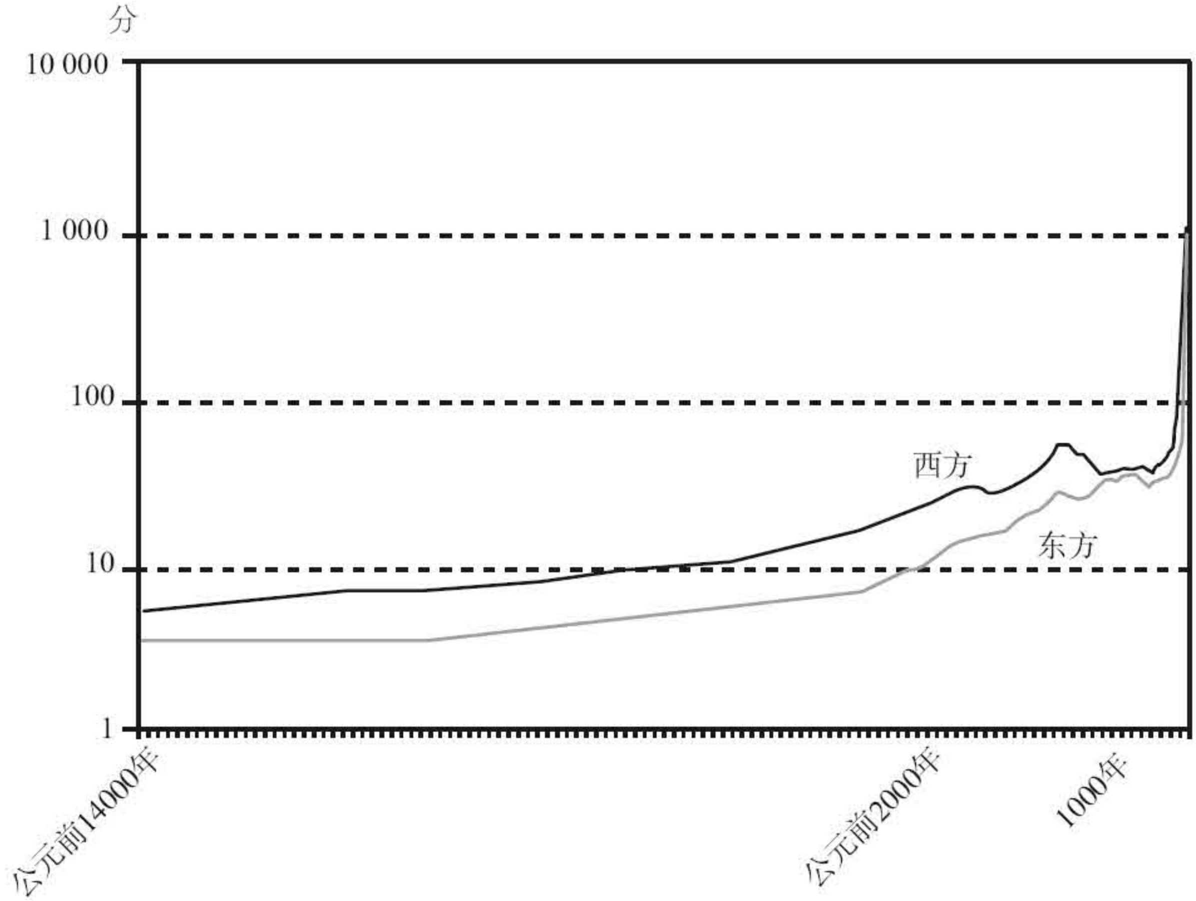
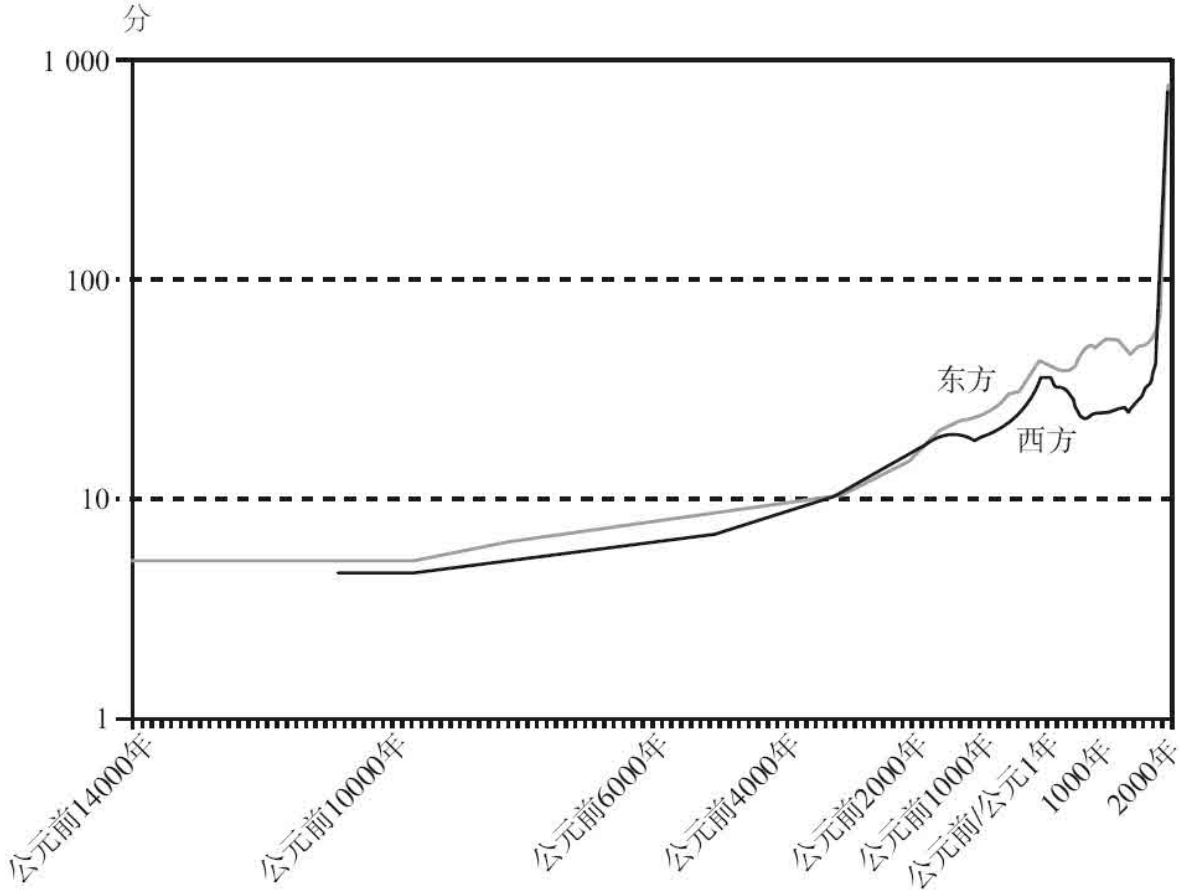


图7.5 公元前14000~公元2000年东方和西方的社会发展分数（以对数–线性标尺标绘，将所有西方的分数降低20%，将所有东方的分数提高20%）



假如图7.4或图7.5是正确的，那么《西方将主宰多久》中的一切就都是错误的。然而，我们可以自信地说，两张图都是错误的。图7.4将西方的分数提高了20%，而将东方的分数降低了20%，告诉我们公元前/公元1年时罗马帝国的发展分数只比1900年工业化时期的日本低5分，这不可能是正确的。图7.5则相反，将东方的分数提高了20%，而将西方的分数降低了20%，意味着东方的发展在商朝之前的时代就比西方在波斯帝国时期快；而西方是直到1828年，鸦片战争前夕时才赶上东方的；并且西方的统治已经结束（实际上是在2003年）。所有这些看来都是不可信的。

因此我的结论是：（1）我估计的误差恐怕不会超过10%，并且肯定不会超过20%；（2）即使误差超过了10%，我试图解释的基本历史模式也仍然是有效的。

指数的直观显示

计算社会发展分数是一回事，展示结果则完全是另一回事。所有可以想象的展现方式都不可避免地会突出信息的某一方面而忽略另一方面。因此，对指数本身以及我在《西方将主宰多久》一书中对其所做的解释的另一种可能的反对意见，或许是我选择的视觉展示方式模糊了其他同样适当的对记录的解释。

感谢斯坦福大学经济系研究生艾萨克·奥珀（Isaac Opper），有一个问题已经引起了我的注意。如我在第二章中所解释的，我计算分数时，给每个特性打的最高分值为250分（均为公元2000年），分数除以250，才能知道什么样的表现可以得到1分，然后我又设定指数上值得记录的最低分值为0.01分。其结果是3项特性的分数一路跌到了0分，分别是社会组织（西方于公元前8000年，东方于公元前4000年）、战争能力（西方于公元前3500年，东方于公元前1750年）、信息技术（西方于公元前3500年，东方于公元前1400年），但能量获取没有。这是因为人们如果不能获取到最低量的能量，即每天4 000卡（得4.36分），他们就无法生存。

结果，无论在东方还是在西方，能量获取分数都在社会发展总分中占到了90%以上，直到公元前100年（见表7.3）；甚至自那以后又过了2 000年，直到20世纪，社会组织、战争能力和信息技术的分数全都有了爆炸式增长以后，能量获取在总分中仍然占到75%以上。

表7.3 社会发展总分中能量获取所占百分比

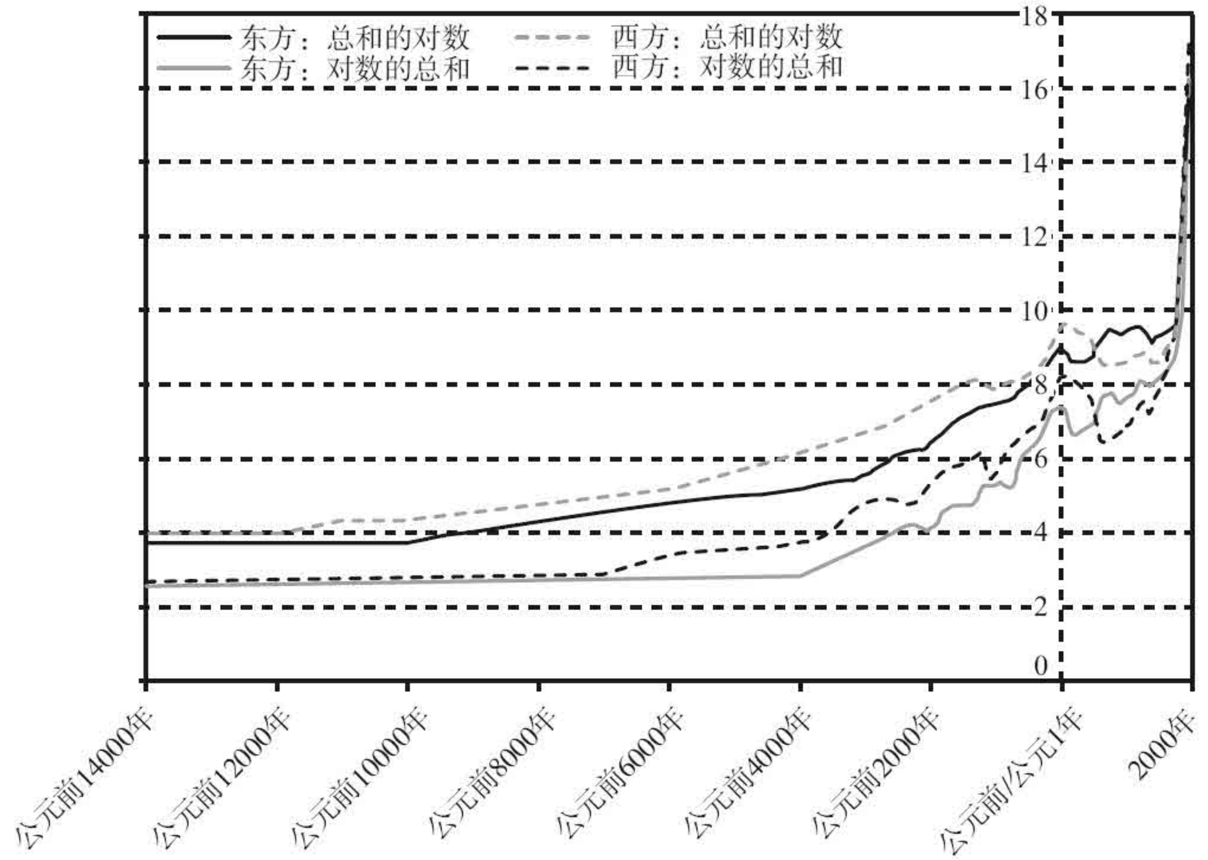
年代	西方（%）	东方（%）	年代	西方（%）	东方（%）
公元前14000~前4000年	100*	100	公元前100年	87	87
			公元前/公元1年	78	86

公元前3500年	92	100*	公元100年	78	88
公元前3000年	97	99	公元200年	78	96
公元前2500年	94	99	公元300年	81	95
公元前2250年	98	99	公元400年	80	94
公元前2000年	97	99	公元500年	88	93
公元前1750年	97	98	公元600年	95	84
公元前1500年	97	98	公元700年	96	76
公元前1400年	97	98	公元800年	94	76
公元前1300年	97	98	公元900年	94	80
公元前1200年	97	97	公元1000年	94	77
公元前1100年	98	97	公元1100年	92	78
公元前1000年	98	98	公元1200年	92	78
公元前900年	98	98	公元1300年	88	81
公元前800年	97	98	公元1400年	96	86
公元前700年	96	97	公元1500年	88	83
公元前600年	95	97	公元1600年	87	83
公元前500年	94	97	公元1700年	85	79
公元前400年	95	96	公元1800年	82	79
公元前300年	95	91	公元1900年	59	75
公元前200年	91	90	公元2000年	28	20

*社会组织的分数，在西方是从公元前8000年，在东方是从公元前3500年开始记录的，但直到公元前3500年的西方，公元前3000年的东方时，它们的分值仍然太小，在社会发展总分中的贡献不到0.5%，意味着它们作为舍入误差消失了。

我在《西方将主宰多久》一书中讨论了能量获取分数是怎样淹没了其他3个特性的，但随意地为其他3个特性加权或提高分数，制造的问题将比解决的问题更多。并非所有人都同意我的观点，但艾萨克·奥珀指出，有更容易的办法让另3个特性更清楚明确。

图7.6 以对数–线性标尺标绘的社会发展分数



*上方的两条线展示了西方和东方4个特性的总和的对数（将4个特性的分数相加，然后计算总和的对数）；下方的两条线展示了西方和东方4个特性的对数的总和（先计算4个特性的分数的对数，然后再将4个对数相加）。对数的总和虽然直观上不够明显，但对社会组织、战争能力和信息技术在前现代时期的小变化更为敏感。
资料来源：艾萨克·奥珀绘制

展现前现代时期社会发展分数的变量的最简单的办法，是在对数－线性图上标绘数据。我在《西方将主宰多久》一书中的某些点上这样做了，将4个特性的分数相加，然后计算其总和的对数。这有其用处，但如果我们分别计算4个特性的对数，再把对数相加，得出一个单一的分数，我们就能绘出一幅不再由能量获取的高分数主宰的图了。图7.6中上方的两条线分别展示了西方和东方特性总和的对数，下方的两条线则分别展示了特性对数的总和。可以看到，将特性的对数相加后产生的曲线，对于20世纪前社会组织、战争能力和信息技术哪怕极小的变化都更加敏感。

公元前2200年后西方核心地带埃及旧王国和阿卡德王国的崩溃，公元前2000年左右东方的陶寺的毁灭，以及公元前771年中国西周王朝的灭亡——当以其他形式绘图时，都被能量获取分数的持续性淹没了——现在却都非常清晰地显示出来了。而且，自20世纪60年代曾被众多的学者极度轻视的西方后罗马时代的瓦解，现在也凸显了。

图7.6 中下方的线也改变了图中左手侧斜线的形状。该图不再单纯由能量获取水平（最重要的是农业的出现）驱动，其对于其他特性更强的敏感性，已将西方领先于东方的时间从约公元前12500年（冰期结束时）推迟到约公元前7000年，最早的人口超过千人的定居点（贝达、巴斯塔、恰塔尔休于等）出现时。现在看来，西方发展的速度在公元前3500年左右最早的国家出现时加快了，然后在公元前300年左右再次加快，当时罗马帝国将地中海沿岸的大部分地区统一在一个单一的政治框架之下，标志着又跨越了一个门槛。不过，在这两个时间之间，公元前2200年和公元前1200年的崩溃，现在以趋势线上突然中断的形式清晰地凸显出来，两次中断后又都迅速地恢复了正常。

东方古代时期的曲线的变化也很有趣。这回农业的出现也没有产生很大影响，发展速度是直到公元前4000年，西坡成长为约1 000人的定居点后，才真正加快的。东方的曲线此后平稳上升（在大约公元前2000年、公元前1600年和公元前800年中断，但都比西方大约公元前2200年和公元前1200年的大崩溃要平和得多），直到大约公元前600年，发展再次提速，直到大约公元100年。进一步的比较研究应当能表明，东方曲线相对于西方的平稳性，是代表着历史的真实，还是只因为我们对西方考古和古代历史知识了解得更详细。

将特性分数的对数相加，而不是计算特性分数总和的对数，对公元1年后2 000年间的曲线形状影响较小，因为到了这个时候，社会组织、战争能力和信息技术的分数已经足够高，无论用哪种办法都可以

显示了。然而，将特性分数的对数相加，的确突出了偏离总体趋势线的程度。

我写《西方将主宰多久》一书的目的之一，是进一步支持中国的发展在明朝和清朝并没有停滞这个观点。图7.6中下方东方的曲线使这一点甚至更加清楚，公元1400~1800年的分数紧紧追随着汉朝之后更长期的趋势。

将特性分数的对数相加，也有助于使西方在公元1400~1800年的爆炸式发展显示得更为清楚，强调了现代西方的主宰并不是以“东方的衰落”为基础的：东方社会在1400~1800年间表现得不错，但西方社会表现得更好。前现代时期全球化对西方社会发展的驱动，像公元前最后3个世纪的地中海化对罗马社会发展的驱动一样有利。

图7.6中下方西方的线突出了我在《西方将主宰多久》一书中所论述的两个论点。第一个是早期现代欧洲的发展的确与罗马共和国的发展有共同之处，而且17~18世纪欧洲人在其社会发展方面重新获得了将近2 000年前的水平时，谈论一场“古代和现代的战争”，是非常正确的；第二个是尽管早期现代的欧洲成绩不俗，但真正革命性的变化还是发生于1800年之后，是随着欧洲西北部的人们开发并利用了化石燃料中的能量而发生的。将特性分数的对数相加，为社会发展提供了一个比计算特性分数总和的对数更灵敏的可视化表现形式，然而能量获取仍然是人类历史的基础。

物质与文化的角力

我曾经说过，社会发展指数的主要贡献之一是迫使分析家分析得更加清楚明确。指数是形容需要做出解释的模式的宝贵工具，那些像我一样持这种观点的人都不得不详细地讲清自己的证据和方法。那些

持不同意见或认为指数的设计或运用不合理的人，也能够清楚地看到指数是怎样运作的，从而进行详细的批驳。争论的双方都不必陷入含糊的斥责。这种含糊的斥责在20世纪八九十年代关于新进化论的争论中曾非常普遍。

在《西方将主宰多久》一书中，我只关注了世界上两个地区，但社会发展指数无疑可以扩大为真正的全球性工具。我常常想，这可以为解决社会科学中一些长久争议不休的问题做些贡献，比如物质力量和文化力量在塑造历史方面哪个相对更重要。

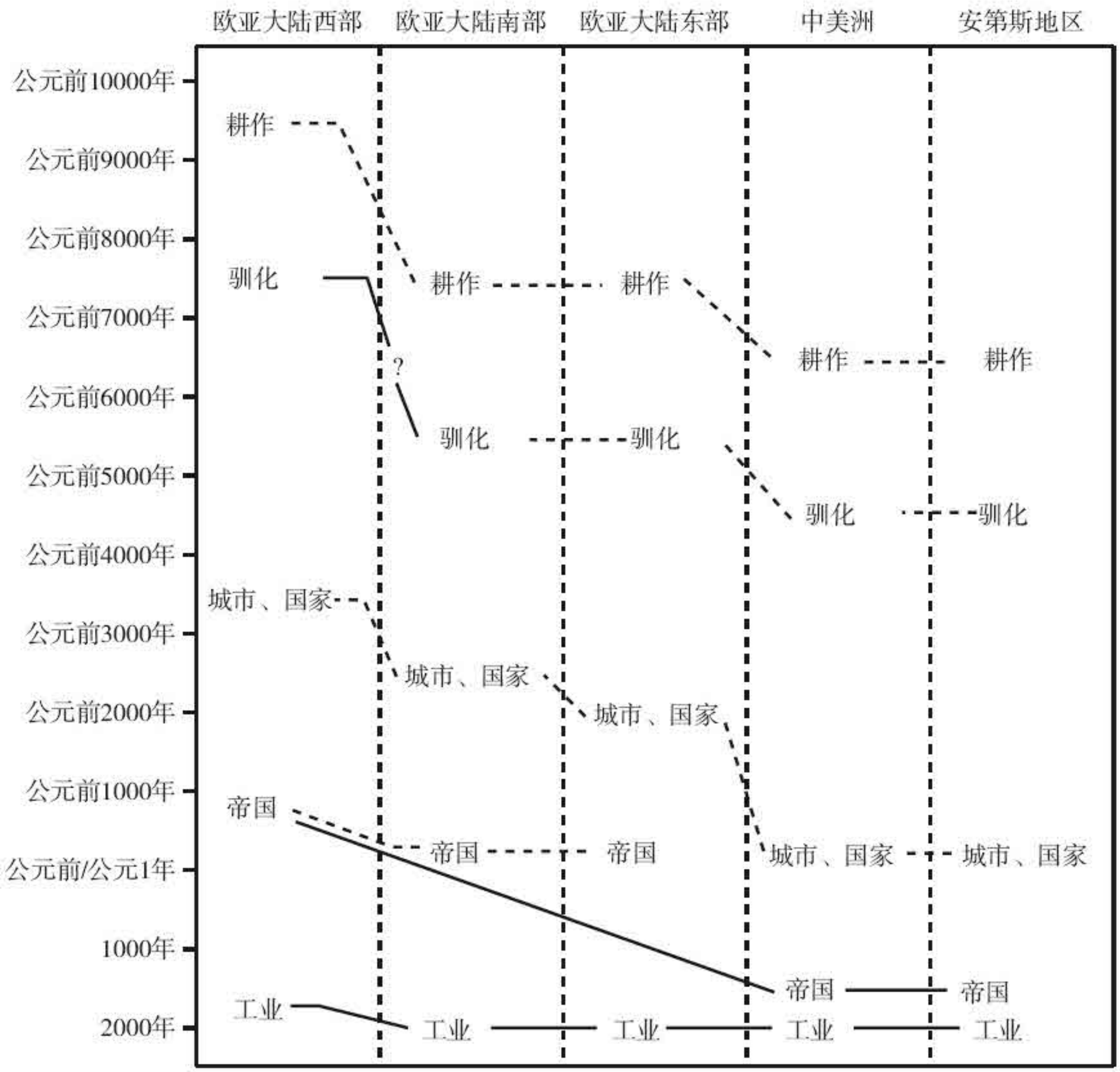
在《西方将主宰多久》一书中，我严厉地批判了唯物主义者，提出过去16 000年来东方和西方社会发展惊人的相似性，表明两个地区的文化特性并没有很大不同。与之一脉相承的是，每个时代都有其需要的思想（或者也许是其值得拥有的思想）。不过，这至多仍是一种假设，而一种适当的全球性社会发展指数，显然将是进一步检验这一观点的工具，将世界的不同部分视为历史的自然实验。

最有用的对比，也许是旧世界和新世界的对比，因为自大约公元前10000~公元1500年，两个时段基本上没有进行有意义的接触。欧亚大陆和大洋洲的澳大利亚也是如此，但新世界还有另一个有利条件，使其与欧亚大陆的对比富有成果。新世界有一处像旧世界而不像澳大利亚的特点，就是在冰期结束时，也有一个包含有大量且集中的可供驯化的动植物的幸运纬度带。澳大利亚的西南部的确也有一个地中海气候区，但在冰期结束时却没有进化出同样丰富的小麦、大麦、稻子、玉米、土豆、绵羊、山羊、猪或牛，这使得驯化动植物并在本地发展出复杂社会比在欧亚大陆困难得多。

如果文化当真是个因变量，我们也许要指望新世界的核心地带以类似于旧世界核心地带的方式扩张，引发同样一系列后发优势、发展悖论、社会崩溃，以及我在《西方将主宰多久》一书中描绘的所有其他现象。

如果将社会发展指数扩展到新世界，就会看到类似事情的确发生了。驯化动植物首先出现在新世界的幸运纬度带，在瓦哈卡地区和秘鲁高原。随着这些地区的社会发展，城市和国家出现了，新的核心地带（即尤卡坦半岛、墨西哥河谷）在最早驯化动植物的地区旁形成了。图7.7是一幅简图，展现了古代复杂社会的5个地区主要的文化转型，加强了人们的这一印象：旧世界和新世界最早的农业核心地带，遵循大致相同的时间表，在很大程度上独立地历经了同样的发展阶段。

图7.7 世界5个地区古代社会文化发展的大舞台（实线代表变革从一个地区到另一个地区的扩散；虚线则非）



耕作（人类对植物生命周期的干预，所产生的选择压力会导致植物出现非自然的大种子）开始与驯化（人类的干预使动植物的基因发生重大改变，使它们转化为只能在人类继续干预下才能生存的新物种）开始的时间间隔，在旧大陆普遍为大约2 000年；在新大陆普遍为大约4 000年，可能是因为新世界的农作物不及旧世界的适应性强。例如，将墨西哥类蜀黍转化为玉米，比将野生小麦、大麦或稻子转化为它们的驯化品种，需要的基因转变要多得多。

然而，由于动植物驯化到城市、国家兴起之间的时间间隔，新世界（约3 000年）比旧世界（3 000~ 4 000年）短，这被部分拉平了。在旧世界，又过了1 500~3 000年，像埃及旧王国和中国商朝这样的国家才变成了真正的帝国，统治着200万以上平方千米的疆域和数千万人口。在新世界，在像莫切文化和特奥蒂瓦坎这样最早的真正国家兴起大约1 500年之后，征服者才来到美洲，切断了本土人的试验；但是在那个时候，印加和阿兹特克的扩张已经产生了与旧世界最早的帝国规模大致相当的组织。假如没有被侵扰，那么再过1 000年，它们完全有可能沿着旧世界帝国道路发展。

作为一名外人看新世界的考古记录，我的印象是：专家根据这份资料算出的社会发展指数，将表明美洲系统性地落后于欧亚大陆，其模式与我在《西方将主宰多久》提出的论点非常相符。在能量获取方面，新世界的一些农作物比旧世界的要强，但美洲缺乏役畜一定严重影响了人均可获得的能量。缺少这样有用的动物，很可能是新世界的运输中车轮使用得非常有限的很大原因，尽管在欧亚大陆，由人推动的独轮车于公元前5世纪在希腊、公元前1世纪在中国，已经显然独立地发明出来了。

在信息技术方面，无论旧世界还是新世界，记录文字和数字的系统都是伴随着最早的城市和国家的出现而投入使用的，然而与大约公元前1500年埃及和美索不达米亚的实践，或者大约公元前500年中国的实践相比，中美洲和安第斯地区在大约公元1500年时对这些技术的使用，似乎仍然非常有限。

在战争能力方面，一些大约是在旧世界最早的国家出现的时代发明的技术（例如堡垒等），在新世界大约也是在同样的时代出现的，但是在随后的1 500年中，这些技术在美洲的传播不及欧亚大陆快。而且一些其他技术（例如青铜武器和盔甲），在美洲根本没有出现；由

于美洲所有有可能进化为驯服马的野马品种在人类到来之后都消失了，新世界的军队无疑从来没有发明过战车和骑兵。

弓箭在新旧世界的命运更是有趣。弓是60 000年前在非洲发明的，然后传遍了旧世界。到公元前第一个千年时，强劲有力的合成弓已经在旧世界所有的复杂社会应用了，十字弩也在中国发明了。然而，据我们至今所知，最早移居美洲的人并没有引进弓，也没有任何迹象表明有人重新发明了弓，直到大约公元前2300年，阿拉斯加和加拿大的“北极小工具传统”遗址才出现了箭头。这些武器随后非常缓慢地在北美传播，直到大约公元1100年才传到中美洲，并且始终没有达到旧世界弓的精致程度。

相反的是，城市规模似乎在新世界的某些部分发展得比旧世界要快。到公元500年时，今墨西哥境内的古代印第安文明特奥蒂瓦坎可能已有10万~20万居民，比欧亚大陆任何地区最早的城市出现后1 000年内的任何遗址都大得多。最早达到特奥蒂瓦坎可能达到的规模下限的旧世界遗址，是大约公元前700年的尼尼微；直到公元前3世纪时才有城市达到其上限，是亚历山大。

对新旧世界的这些差异的解释，可能已经具备。贾雷德·戴蒙德在其《枪炮、病菌和钢铁：人类社会的命运》一书中，指出欧亚大陆比之美洲有三大地理优势，可能在很大程度上是使其在我所谓的社会发展中领先的原因。

戴蒙德说，首先，在冰期结束时，欧亚大陆拥有比美洲更为丰富的自然资源基础，使得亚洲西南部和东部的人们比中美洲或安第斯地区的人们更容易驯化植物和动物。其次，新世界有可能驯化的大型（即重达100磅以上的）哺乳动物，在最早的人类殖民该大陆后很快出现的巨型动物灭绝现象中大多被消灭了。最后，甚至大陆的布局都对美洲原住民不利。欧亚大陆基本上是东西走向的，因而起源于西南亚的观念、制度和实践，能够在地理学家们称之为“生物群系”的统一的

生态区中传播数千英里，到达欧洲或中国。美洲则相反，基本上是南北走向的，意味着在中美洲或安第斯地区崭露头角的观念、制度和实践，只能在小股人群（相对于旧世界而言）中流传，然后才能跨越极其不同的生物群落区。因此，新世界的观念、制度和实践要很长时间才能浮现，要更长时间才能传播。

由于大陆数量太少，戴蒙德的观点很难验证，尽管最近一项对语言多样性的研究的确提供了至少是有限的支持。新世界的社会无疑有其独特的特色，它们有早熟的（相对于旧世界社会而言）城市化，却又较慢采用新的作战方式和信息技术，这一现象仍亟待解释。将社会发展指数延伸到全球尺度上，将更容易看出戴蒙德的地理框架是否能解释新旧世界的差异，或者我们是否需要赋予文化因素更重要的地位。

直线发展和多线发展的进化

“毫无疑问，直线发展的进化与多线发展的进化之间的争执，是文化进化论方面最令人烦恼的问题之一。”罗伯特·卡内罗在结束他对这个领域的研究时总结道。

不过，像许多这样的问题一样，烦恼很大程度上来自不清楚问题是什么，以及对适度水平的抽象的隐含假设。在一些学者看来，问社会发展是否是直线发展的，似乎意味着问现代化是否只有一条道路（而且，相应的还要问现代化是否不止一种）；而在另一些学者看来，这似乎意味着问更普遍的问题：社会是否会向多个方向发展（相应的，问是否有多重道路，以致谈及发展就是个错误）。

社会发展指数认为，对第一种问题——现代化是否只有一条道路——的回答，是毫不含糊的“是”。指数显示的不仅是从采集者到农民

再到工厂工人并继续发展的一系列非常清晰的进步，也表明了在每种广义的组织下的发展所能达到的一系列像“硬天花板”一样的上限。在指数上，任何狩猎－采集社会都不可能发展到超越6~7分的程度；任何农业村落社会都不可能超越10~12分；任何农业帝国都不可能超越45分。任何狩猎－采集社会或农业村落社会都不可能跨越农业帝国阶段而直接实现工业化——除非受到另一个已经经历了这些阶段的社会的影响；没有任何社会能从草原游牧阶段直接进入工业社会，除非先被一个农业帝国征服。

人类学家、历史学家和社会学家一直无休无止地争论现代化到底意味着什么，但是所有他们通常视为“现代”的社会，都经历了上一段描述中的一两条道路。无论在什么情况下，现代化都包含能量获取激增的意味。这种激增是由借助了化石燃料的威力的工业革命带来的，而其后果又是将能量应用于重新划分的各行各业。关于现代化是多重的还是单一的的争论，很大程度上可归结为对概括这些社会的最有益水平方面的分歧。

如果我们从第二种意义上来看直线发展与多线发展之争问题——社会是否能向多个方向发展——答案在很大程度上取决于我们的时序观点，以及我们将调查的终点设在哪里。如果我们从观察1 000年前的世界开始，我们的回答将是毫不含糊的“是”：社会非常显然是在向多个方向发展的，从卡拉哈里沙漠的猎手到突厥游牧者，从密西西比河畔的农夫到中国宋朝。

然而，在支持多线进化的论点中发挥着巨大作用的文化，在解释这种多样性时，作用就很有限了。在一个给定的生物群落区内，众多社会倾向于以大体相同的方式发展。农业是在幸运纬度带内生根的，也是从那里向外推广的。居住在大草原上的牧民，经历了徒步、乘车和骑马的发展阶段。狩猎和采集则主宰着不支持农业或畜牧业的环境。

又一次，我们的工作所停留的抽象层次也决定着我们能看见什么。譬如说，中国人、印度人、阿拉伯人、基督徒和托尔特克人的生活方式，无疑在各种各样的方面都大相径庭，学者们完全可以随意选择强调这些差异而忽视其共性。然而，社会发展指数迫使我们认识到，生物群落区内的共性是真实存在且至关重要的。决定多线进化的主要因素是地理，而非文化。

这就是说，如果我们只观察21世纪的世界，我们会得出相反的结论，对上述问题的回答将是“否”：社会并不是向多个方向发展的。几千年来，在不同的生物群落区内，地理推动着发展沿不同道路前进，但是随着世界上的幸运纬度带内社会发展的增长，最发达的社会在空间上不断扩张。到公元200年时，欧亚大陆上农业帝国的商人们将商品从旧世界的一端移向另一端；到公元600年时，我们已掌握有个人来往于欧洲到中国全程的确凿证据；到公元1400年时，欧亚大陆上的人已经有了能够可靠地渡过任何海洋的船只；而到公元1900年时，欧洲人及其海外殖民者已几乎将整个世界聚合成一个单一经济体。全球化历程已经发展了几个世纪，一旦有一个社会跨越了化石燃料的门槛，整个世界迅速变成现代社会就是不可避免的了。

于是，对于社会是否会向多个方向发展的问题，我们是回答“是”还是“否”，完全取决于我们所着眼的年代。研究史前时期的考古学家和研究前现代时期的历史学家可以心安理得地回答“是”，但是任何将21世纪包括在内的思想，都迫使我们回答“否”。

我相信，社会发展指数指向合格的直线发展。第一，只有一条道路通向现代化；第二，只有居住在特定地方的人们能走上这条道路；第三，一旦有足够多的人在这条道路上走得足够远了，其他所有人就都会被拽上这条道路。

社会进化的方向

生物进化经常被描绘为一个无方向的过程。博物学家斯蒂芬·杰·古尔德（Stephen Jay Gould）曾以一幅清晰生动的画面，向我们描述了假如我们有什么办法重放生活磁带的话，那么它将极可能不会再导向我们这里了。他说："神圣的磁带播放机掌握上百万个脚本，每个脚本都极其合理。在开端处即使不经意间一个小小的偏转，都会释放出像众多小瀑布一般的结果，形成一个在回顾时会觉得不可避免的特定未来。然而在开始时哪怕最轻的一碰，假如接上了另一个槽口，历史就会转入另一个看似合理的渠道……因此，对我们来说，我想我们只能大喊：噢，神勇的——也是不大可能的——新世界，让这样的人进去吧！"

在进化论是思考人类社会的正确框架的情况下，我们应当像古尔德一样看到生物学和文化之间的连续性，也许得出这样的结论：没有任何事情使得过去15 000年不断上升的社会发展成为不可避免的模式。指数上的数字开始攀升，是因为公元前12700年后全球气候变暖，导致上升的发展成为一种成功的适应。改变了环境背景，你也就改变了适应景观；在一种环境下兴盛，比如发展水平高的群体，到了另一种环境下就不一定兴盛了。

这种观点的隐含之意，似乎是社会进化像生物进化一样没有方向；一种社会发展指数，无论其在了解历史方面的描述性价值如何，都没有预测能力。但这样看也许会将婴儿与洗澡水一起泼掉。生命历史上其实有非常明显、非常强大的模式，很多生物学家都曾得出结论：动物进化到像人类一样聪明——也许像人类一样看待事物——是完全可能的，以致我们也可以称之为不可避免的。

古尔德本人愿意承认，尽管他看到了生命彩票本质的随意性，但贯穿生命历史，有一种"复杂性的右倾"一直在非常稳定地运行着。社会历史尽管时常被像物种历史所受到的同样众多的冲击所打断，然而

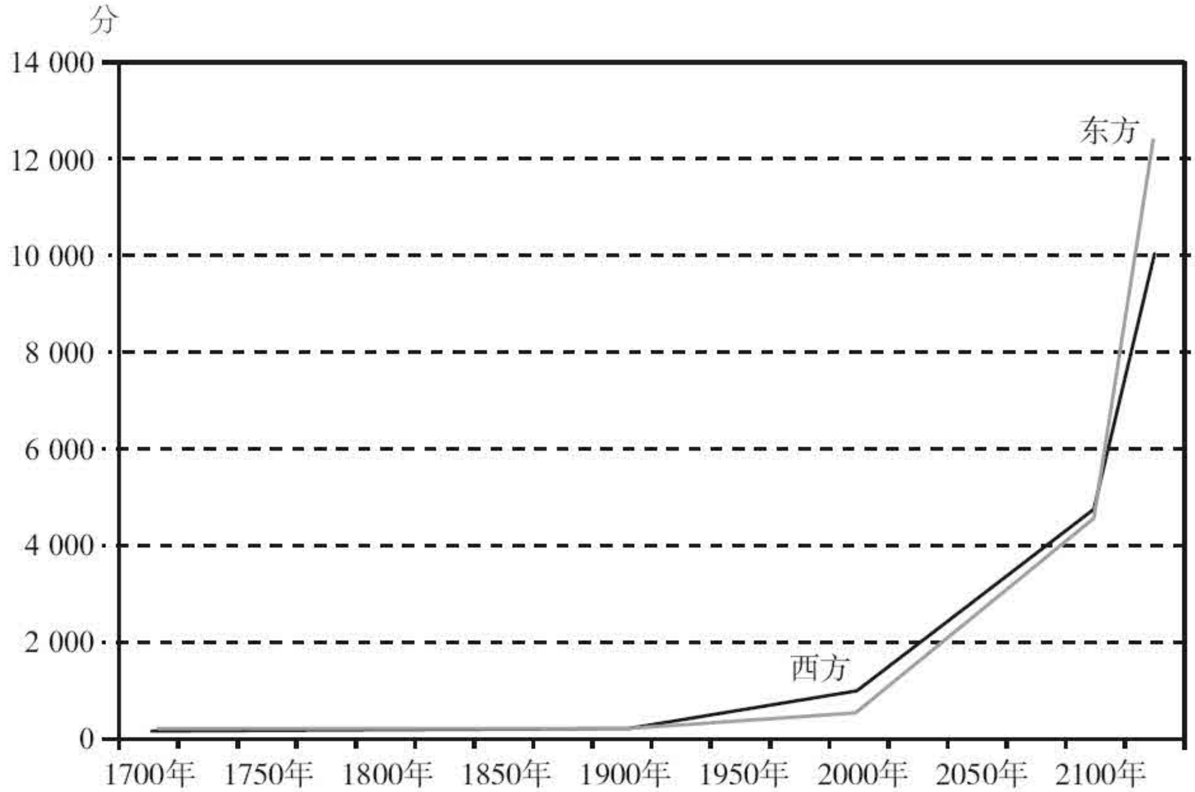
在过去15 000年中，也仍然能看到同样稳定的“右倾趋势”。整个历史明显的含义是——在其他条件都相同的情况下，从长远来看——我们也许能预期社会发展分数会不断地增加。

在《西方将主宰多久》一书中，我提出事实的确如此，并认为如果我们预测社会发展指数在整个21世纪及以后的趋势，我们是可以有一些看法的。我的预测非常粗糙，只是推断东方和西方的社会发展指数在21世纪将以与20世纪相同的速度继续增长。假如情况的确如此，东方的分数将在2103年赶上西方，届时两个地区在指数上都将达到大约5 000分（见图7.8）。

这个论点比本书的任何其他部分都吸引了更广泛的关注。21世纪的发展道路会遵循这样一个过分简单化的直线性预测，是非常不可能的，但是预测的价值不仅在于它能否实现（实际上所有预测都不能实现），更多的是在于迫使我们提问，我们需要将设想改变多少，才能产生彻底不同的结果。

预测20世纪指数趋势线到21世纪的走向时，会认为过去50年“东方的崛起”并非西方主宰的暂时中断。这是中期历史趋势的结果，是由可回溯到一个多世纪前的地理意味的变化（最重要的是，太平洋实际上变小了）驱动的。西方国家政府也许能操纵财富和实力向东方的转移，最老练的西方政策预言家们关注的恰恰是这一点。然而，阻止或逆转东方的崛起，要求有至今尚未确定的会使地理再度强烈改变其意味的发展，并且这些发展是有利于西方复兴的优势的。

图7.8 未来的趋势？对21世纪东方和西方社会发展分数的预测



不过，预测20世纪趋势线的进一步走向的最意味深长的含义，是社会发展在下一个百年将达到5 000分。从公元前14000~公元2000年，发展提高了900分。这是人类从旧石器时代的岩洞壁画发展到互联网而获得的。不过，按照图7.8的设想，2000~2100年发展还将再提高4 000分；如果有什么不同的话，那就是这个分数可能被低估了。新世纪开端这十几年的所有迹象都表明，发展是以指数方式增长的，而不仅是像20世纪末时那样的线性延伸。

又一次，重要的问题不是问我在《西方将主宰多久》一书中讨论的具体预测能否实现，而是需要发生什么情况，才能使现实远离我的预测所依托的设想，使社会发展分数在2100年时远远达不到5 000分。从过去15 000年的社会发展曲线来判断，最有可能的答案似乎是发生一场新的社会崩溃。

5 000分的设想认为人类社会永远摆脱了马尔萨斯理论的限制，但对冰期后的历史显然也有另一种解释：工业化只是将这些限制向外推了推。它推动力度的确很大，使全世界的人口增长了一个数量级，也使数十亿人脱了贫；但对指数的这种解读仍然要提出，工业社会也将面对内在的“硬天花板”，就像曾限制了农业社会发展的“硬天花板”。

罗马帝国和宋朝等农业帝国在社会发展分数超过了40分以后，便开始动摇和衰退，现代社会在1 000~5 000分的某处，也许也会遇到新的“硬天花板”。假如一个拥有100亿人口的世界发展停滞了，或者出现了难以预料的气候变化，或者发生了核扩散，再或者出现了机器人战争、电子战争和纳米战争等分布不平衡的发展，其后果会比农业时代的发展停滞还要恐怖。

指数暗示接下来的50年将是历史上最重要的一段时期。一方面，假如19世纪和20世纪能量增长带来的好运只是一锤子买卖，那么21世纪就将注定成为对所有人来说都是最糟糕的时代；另一方面，假如工业革命被证明只是一个更长时期的能量革命的第一阶段，那么当前这个世纪将肯定会使人类变得面目全非。到21世纪60年代时，也许我们将看清世界走向什么道路。

社会发展指数当然无法解决这些问题中的任何一个，但它可能是确定其中一些问题的宝贵工具。理查德·道金斯（Richard Dawkins）在他的经典著作《自私的基因》（*The Selfish Gene*）的开头推断：“假如有更高级的生物从太空造访地球，为了估量我们的文明水平，他们问的第一个问题会是：‘他们发现进化论了没有？’”假如这样的事情当真发生了，我想他们问的既包括生物进化，也包括社会进化。

致谢

《文明的度量》与我此前所写的《西方将主宰多久》是姊妹篇，但是一本截然不同的书。在《西方将主宰多久》中，我试图讲述最近15 000年来社会发展的故事；而在本书中，我想介绍我在建构那些故事背后的社会发展指数时所使用的论据和论证方法。

和许多书一样，本书也是持续多年的讨论的产物。早在20世纪80年代初，即我在英国剑桥大学读研究生时，就有人向我介绍了社会进化的观点，自那以后，我就不时思考并与他人讨论这个问题。在此期间，我欠下了很多人的人情，特别想感谢以下各位：达龙·阿西莫格鲁（Daron Acemoglu）、詹姆斯·安德森（James Anderson）、约翰·贝内特（John Bennet）、弗兰切斯卡·约翰·谢里（Francesca John Cherry）、埃里克·齐恩斯基（Eric Chinski）、大卫·克里斯汀（David Christian）、杰克·戴维斯（Jack Davis）、斯特凡·德·施皮格利埃（Stephan de Spiegeliere）、贾雷德·戴蒙德、阿尔·迪安、汤姆·加伦特（Tom Gallant）、彼得·加恩西（Peter Garnsey）、班宁·加特（Banning Garrett）、杰克·戈德斯通（Jack Goldstone）、黛博拉·戈登（Deborah Gordon）、史蒂夫·哈伯（Steve Haber）、约翰·霍尔顿、保罗·霍尔斯特德（Paul Halstead）、伊恩·霍德（Ian Hodder）、徐心眉（Agnes Hsu）、帕拉格·康纳（Parag Khanna）、卡拉·克尔凯郭尔（Karla Kierkegaard）、克里斯蒂安·克里斯蒂安森（Kristian Kristiansen）、大卫·莱廷（David Laitin）、迈克尔·莱西希（Michael Lässig）、鲁威仪、安东尼·林（Anthony Ling）、刘莉、安格斯·麦迪森、阿莱西奥·马格纳瓦卡（Alessio Magnavacca）、保罗·马拉尼马、乔·曼宁（Joe Manning）、迈克尔·麦考密克（Michael McCormick）、汤姆·麦克莱伦（Tom McLellan）、乔尔·莫基尔（Joel Mokyr）、苏雷什·奈杜

（Suresh Naidu）、雷维尔·内茨（Reviel Netz）、道格·诺思（Doug North）、乔希·奥伯（Josh Ober）、艾萨克·奥珀、安妮·波特（Anne Porter）、迈克尔·普特（Michael Puett）、库马尔·罗摩克里希纳（Kumar Ramakrishna）、安娜·拉泽托（Anna Razeto）、科林·伦弗鲁（Colin Renfrew）、吉姆·罗宾逊（Jim Robinson）、理查德·萨勒（Richard Saller）、沃尔特·沙伊德尔、格伦·施瓦茨（Glenn Schwartz）、雨果·斯科特－加尔（Hugo Scott-Gall）、史蒂夫·香农（Steve Shennan）、丹·斯梅尔（Dan Smail）、瓦茨拉夫·斯米尔、拉里·史密斯（Larry Smith）、迈克·史密斯（Mike Smith）、安东尼·斯诺德格拉斯（Anthony Snodgrass）、彼得·特明（Peter Temin）、尼克·托马斯（Nick Thomas）、彼得·图尔钦、巴里·温格斯特（Barry Weingast）、托德·怀特洛（Todd Whitelaw）、詹姆斯·惠特利（James Whitley）、格雷格·伍尔夫（Greg Woolf）和诺姆·约菲（Norm Yoffee）。他们都曾帮助我从不同的角度看问题，希望他们能同意我很好地采纳了他们的意见。

如果没有下面这些人，我也根本完不成这本书。感谢普林斯顿大学出版社的罗布·滕皮奥（Rob Tempio）的鼓励；感谢Profile Books图书公司的丹尼尔·克鲁（Daniel Crewe），是他看出原本只被我视作一个数据库的东西可以扩展为一本书；感谢桑迪·戴克斯特拉（Sandy Dijkstra）和阿拉贝拉·斯坦（Arabella Stein）的指教，是他们把大家召集在一起；感谢凯茜·圣约翰（Kathy St. John）的支持和耐心；感谢我父亲诺埃尔·莫里斯（Noel Morris）给我树立的榜样和早年对我的教育，事实证明，他教我的数学派上了用场。